

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/156496>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.



# Towards the Harmonisation of Outcome Measures for Children with **Mitochondrial Disease**

Saskia Koene

## Towards the Harmonisation of Outcome Measures for Children with Mitochondrial Disease



**ZonMw**

**Radboudumc**

**Radboud University Nijmegen**



The research presented in this thesis was performed at the Nijmegen Centre for Mitochondrial Disorders (currently: Radboud Center for Mitochondrial Medicine) at the department of Paediatrics of the Radboudumc, Nijmegen, the Netherlands and was additionally financially supported by the Netherlands Organisation for Health research and Development (ZonMw), Stofwisselkracht, Stichting Energy4all, Zeldzame Ziekten Fonds and the Tjallingh Roorda Foundation. Printing of this thesis was financially supported by the Radboud University, Nijmegen, the Netherlands and Sigma-Tau HealthScience International B.V. All funding is gratefully acknowledged.

**ISBN13/EAN:** 978-90-9029553-4

**Design** Studio ImVorm, Nijmegen

**Print** Ipskamp Drukkers Nijmegen

© 2016, Saskia Koene, Nijmegen, the Netherlands

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronically, mechanically, by photo-copying, recording or otherwise, without the prior written permission of the holder of the copyright. The copyright of the articles that have been published has been transferred to the respective journals. Copyright of published work: Nature Publishing Group (Chapter 1), Elsevier (Chapter 2), Wiley-Blackwell (Chapter 3), Springer Publishing (Chapter 4, 5 and 11) and Wolters Kluwer (Chapter 6).



# Towards the Harmonisation of Outcome Measures for Children with Mitochondrial Disease

## **Proefschrift**

ter verkrijging van de graad van doctor aan de Radboud Universiteit Nijmegen  
op gezag van de rector magnificus, volgens besluit van het college van decanen  
in het openbaar te verdedigen op donderdag 28 april 2016 om 12.30 uur precies.

door

Saskia Koene  
geboren op 19 mei 1984  
te Weert

### Promotor

Prof. Jan A.M. Smeitink

### Copromotor

Dr. Imelda J.M. de Groot

### Manuscriptcommissie

Prof. Frans G. Russel

Prof. Joost P.H. Drenth

Prof. Patrick F. Chinnery

*University of Cambridge, UK*

### Paranimfen

Erik Koene

Laura S. Storm-Schulz

voor mijn ouders



## Table of Contents

Prologue	Introduction	16
	Problem definition	36
Scope, aims and outline of this thesis		50
	Scope	52
	Aims	53
	Outline	54
PART I: STUDYING INTERVENTIONS IN MITOCHONDRIAL DISORDERS: PAST, PRESENT AND FUTURE		56
Chapter 1.	<b>New treatments for mitochondrial disease - no time to drop our standards</b>	60
	<i>Nature Reviews Neurology 2013 Aug;9(8):474-81</i>	
Chapter 2.	<b>Developing outcome measures for paediatric mitochondrial disorders which complaints and limitations are most burdensome to patients and their parents?</b>	82
	<i>Mitochondrion 2013 Jan;13(1):15-24</i>	
Chapter 3.	<b>Towards the harmonisation of outcome measures in children with mitochondrial disorders</b>	108
	<i>Developmental Medicine and Child Neurology 2013 Aug;55(8):698-706</i>	
PART II: NATURAL DISEASE COURSE		134
Chapter 4.	<b>Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: what we learned from 130 cases</b>	138
	<i>Journal of Inherited Metabolic Disease 2012 Sep;35(5):737-47</i>	
Chapter 5.	<b>Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A&gt;G mutation</b>	164
	<i>Journal of Inherited Metabolic Disease Reports 2012 Nov;35(6):1059-69</i>	



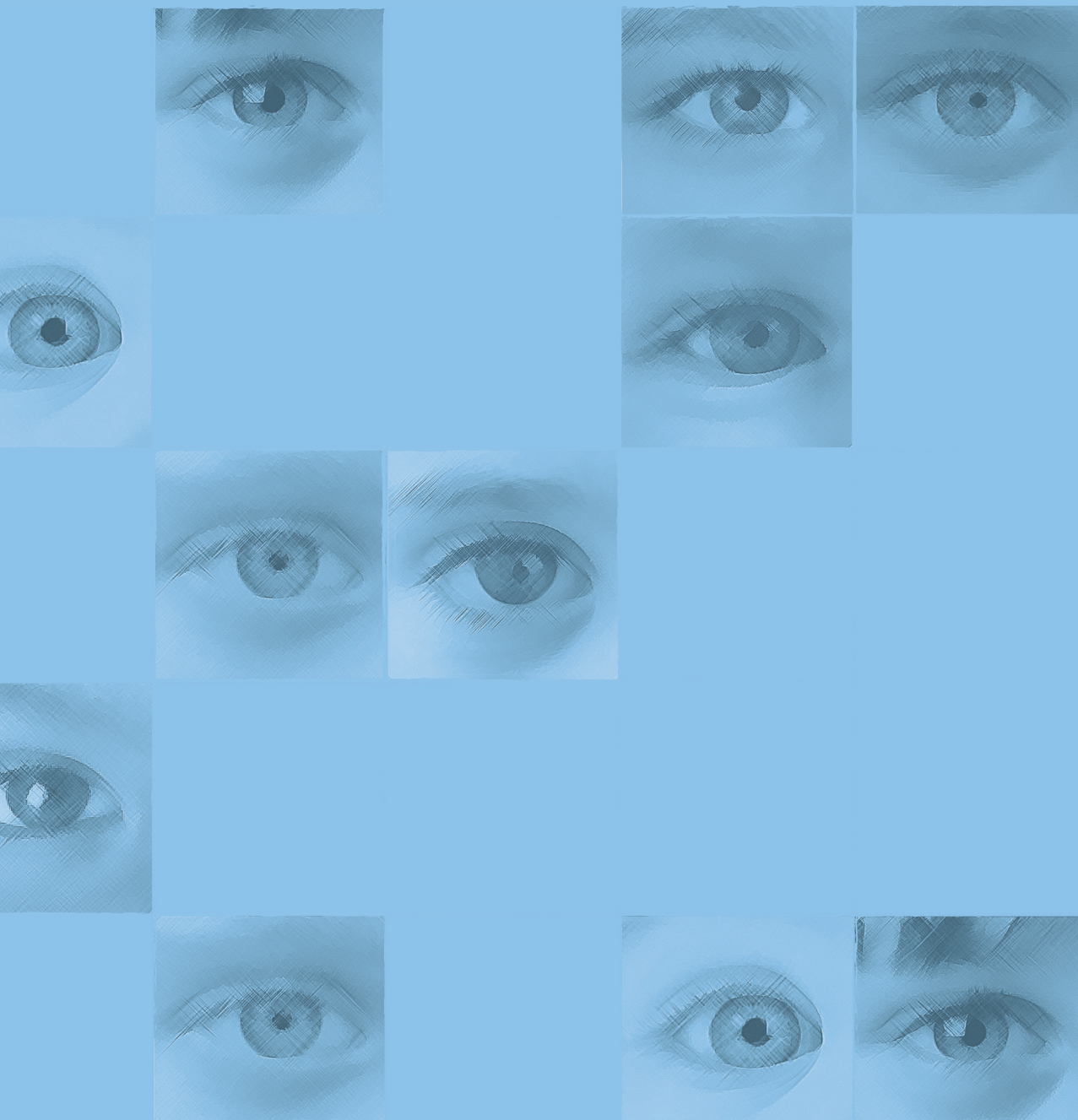
PART III: VALIDATION STUDIES		186
Chapter 6.	<b>The value of using 3D accelerometry in estimating daily physical activity in children with mitochondrial disease</b> <i>Submitted</i>	190
Chapter 7.	<b>The International Paediatric Mitochondrial Disease Scale</b> <i>Submitted</i>	216
Chapter 8.	<b>The assisted 6-minute cycling test: an exploratory study in children</b> <i>Muscle &amp; Nerve; accepted manuscript online Dec 30 2015</i>	238
Chapter 9.	<b>Is 2D speckle tracking echocardiography useful for detecting and monitoring myocardial dysfunction in adult m.3243A&gt;G carriers? - a retrospective pilot study</b> <i>Submitted</i>	260
Chapter 10.	<b>Serum Fibroblast Growth Factor 21 (FGF21) levels in adult m.3243A&gt;G carriers: clinical implications</b> <i>Neurology 2014 Jul 8;83(2):125-33</i>	284
Chapter 11.	<b>Serum GDF15 levels correlate to mitochondrial disease severity and myocardial strain, but not to disease progression in adult m.3243A&gt;G carriers</b> <i>Journal of Inherited Metabolic Disease Reports 2015;24:69-81</i>	308
Supplementary Documents		332
Discussion	General discussion <i>Partially in press by EMBO Molecular Medicine</i> Future prospects	406 422
Summaries	Summary Samenvatting Dankwoord	440 448 456
List of Publications & Curriculum vitae	List of Publications Curriculum vitae	464 470

## List of abbreviations

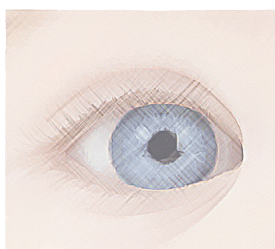
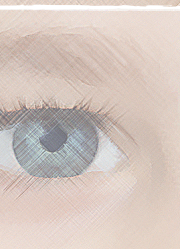
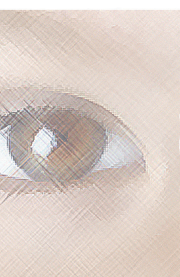
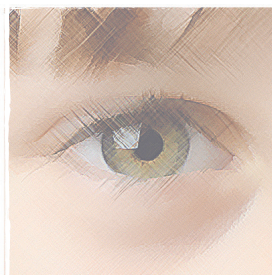
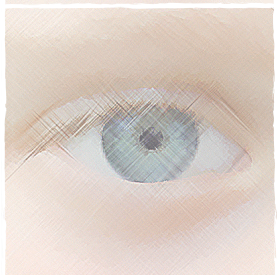
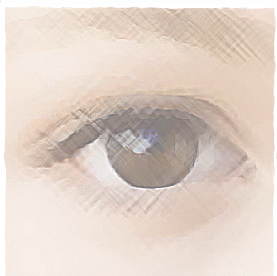
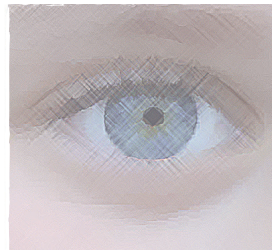
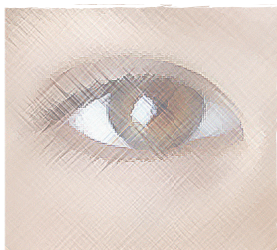
2DSTE	2-dimensional speckle tracking echocardiography
3D	3-dimensional
4-CV	4 chamber view
6MWT	6-minute walking test
95%CI	95% confidence interval
A6MCT	assisted 6-minute cycling test
ADL	Activities of daily life
AMPK	5' adenosine monophosphate-activated protein kinase
ANT	Adenine nucleotide transporter
ATP	Adenosine triphosphate
$\beta$	Regression coefficient
BMI	Body mass index
C	Complex
CI	Confidence interval
CoQ10	Co-enzyme Q10
CM	Congenital myopathy
cMRI	cardiac magnetic resonance Imaging
CNS	Central nervous system
CP	Cerebral palsy
CPEO	Chronic progressive external ophthalmoplegia
CSF	Cerebrospinal fluid
CV	Covariance
DM	Diabetes mellitus
DMD	Duchenne muscular dystrophy
dNTPs	deoxynucleotide triphosphates
EF	Ejection Fraction
ELISA	Enzyme-linked immuno sorbent assay
EMA	European Medicines Agency
FAH	Flavin adenine dinucleotide
FADH	Flavin adenine dinucleotide, reduced form
FDA	Food and Drug Administration
FGF21	Fibroblast growth factor 21
FRDA	Friedreich's ataxia
FS	Fractional Shortening
FSGS	Focal segmental glomerulosclerosis
GCS	Global circumferential strain
GDF15	Growth and differentiation factor 15
GFR	Glomerular filtration rate
GLS	Global longitudinal strain
GMFM	Gross Motor Function Measure
GRS	Global radial strain
ICC	Intra-class correlation coefficient
ICF-CY	International Classification of Functioning, Disability and Health for Children and Youth
IPMDS	International Paediatric Mitochondrial Disease Scale
IQR	Interquartile range



LGMD	Limb girdle muscular dystrophy
LHON	Leber's hereditary optic neuropathy
m.3243A>G	A to G nucleotide change at position 3243 in mitochondrial DNA
MD	Myotonic dystrophy
MECD	Multiple enzyme complex deficiencies
MELAS	Mitochondrial encephalopathy, lactic acidosis and stroke-like episodes
MERRF	Myoclonus epilepsy with ragged red fibres
MFM	Motor Function Measure
MIDD	Maternally inherited diabetes and deafness
mtDNA	Mitochondrial DNA
MRI	Magnetic resonance imaging
MRC	Medical Research Council
MRS	Magnetic resonance spectroscopy
<i>n</i>	Number of observations
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide, reduced form
NMD	Neuromuscular disorders
NMDAS	Newcastle Mitochondrial Disease Adult Score
NPMDS	Newcastle Paediatric Mitochondrial Disease Scale
OXPHOS	Oxidative phosphorylation
<i>p</i>	Level of significance
PDHC	Pyruvate dehydrogenase complex
PEDI	Pediatric Evaluation of Disability Inventory
PGC	Peroxisome proliferator-activated receptor gamma coactivators
PGD	Pre-implantation genetic diagnostics
PPAR $\gamma$	Peroxisome proliferator-activated receptor gamma
PPi	Pyrophosphate
QoL	Quality of life
QoLP	Quality of physical health
QoLM	Quality of mental health
<i>r</i>	Correlation coefficient
rev	Revolutions
ROS	Reactive oxygen species
SD	Standard Deviation
SIRT1	NAD-dependent protein deacetylase sirtuin 1
SMA	Spinal muscular atrophy
TCA	Tricarboxylic acid
UEC	Urinary epithelial cells
WHO	World Health Organisation



# Prologue



# Introduction



## Introduction into Mitochondria and Mitochondrial Disease

Mitochondrial disorders are the most common cause of inherited metabolic disease, with an estimated prevalence of 1 in 5,000 to 1 in 8,500 live births.<sup>1,2</sup> The genetic and consequently the biochemical background of these diseases varies widely, as do the clinical signs and symptoms.<sup>3</sup> The clinical spectrum varies from neonatal fatal lactic acidosis, to paediatric multi-system diseases and adult oligo-system disorders, to dormant (asymptomatic) carriers. To date, no cure is available, but new substances aimed at providing a less poor disease outcome are under development.<sup>4</sup> To judge the effectiveness of these compounds, clinical studies need to be executed. Previously, a Cochrane review on treatment studies in patients with mitochondrial disease stated that the quality of many of the previously performed studies is low.<sup>5</sup> One of the main problems arising in these low quality studies includes the heterogeneity of the, often small-numbered, study population and the lack of sufficiently sensitive end points.

This thesis aims to select and validate both clinical and laboratory outcome measures for mitochondrial disease in children. In this introductory chapter, we illustrate the heterogeneity and complexity of mitochondrial medicine by providing information on the molecular background underlying mitochondrial disease and describing the heterogeneous phenotype. To illustrate the infancy of this field in clinical trial methodology, we will provide a short history on treatment studies in mitochondrial disease.

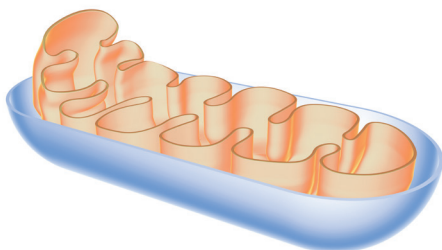
## The Heterogeneity of Mitochondrial Disease

Three main characteristics of mitochondrial medicine are responsible for the complexity and heterogeneity of mitochondrial disease. First of all, mutations in any of the ~1,100 proteins involved in mitochondrial assembly, structure, function and maintenance may be involved in the pathogenesis of a mitochondrial disease. Furthermore, in case of a mutation in the mitochondrial DNA, the percentage in which the mutation is present (*i.e.* heteroplasmy) may vary between tissues or even within tissues. Finally, although mitochondria are present in all cells except for the erythrocytes, not all tissues seem to be equally sensitive to the failing mitochondria which can only partly explained by differences in tissue distribution of affected proteins.

### Mitochondrial structure and function

Mitochondria are found in almost every cell in a varying number per cell type. For example, a human liver cell may contain 1,000-2,000 mitochondria per cell; whereas a mature erythrocyte contains none.<sup>6</sup> Mitochondria exist, at least in dividing cells in culture and in neurons, in a dynamic mitochondrial network, showing constant fusion and fission.<sup>7</sup>

Mitochondria contain outer and inner membranes that are both composed of a phospholipid bilayer and proteins (Figure 1). The outer mitochondrial membrane encloses the entire organelle and has a phospholipid:protein ratio of 1:1. It contains protein import machineries, protein channels that allow diffusion of small molecules into the mitochondrial inter-membranous space. The inner mitochondrial membrane has a very high (4:1) protein to phospholipid ratio and contains numerous proteins such as the respiratory chain complexes, adenosine triphosphate (ATP) synthase and protein and metabolite import machineries.<sup>8</sup> The inner mitochondrial membrane is compartmentalised into numerous cristae, which greatly expand the surface area of the inner mitochondrial membrane, increasing the number of energy generating proteins and thereby the ability to produce ATP.<sup>9</sup> This characteristic architecture of the mitochondrial inner membrane is thought to be maintained by the interaction of the mitochondrial contact site and cristae organising system which is closely related to both inner and outer mitochondrial membrane.<sup>10</sup> The mitochondrial matrix is the space enclosed by the inner mitochondrial membrane. The matrix contains a highly-concentrated mixture of hundreds of enzymes, mitochondrial ribosomes, transfer (t)RNAs and several copies of mitochondrial DNA (mtDNA),<sup>11</sup> which are double-stranded circular DNAs.



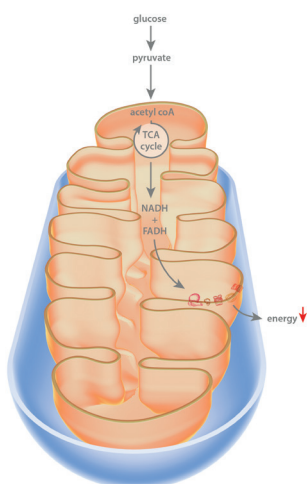
**Figure 1.**  
**Artist impression of a mitochondrion.**

Mitochondria are involved in several metabolic processes such as haem synthesis, heat production, the storage and signalling of calcium ions, programmed cell death and steroid synthesis.<sup>12</sup> The most prominent role of mitochondria, in the context of this thesis, however is to provide the cell with energy in the form of ATP.<sup>11</sup>

ATP production from glucose occurs via glycolysis, the tricarboxylic acid (TCA) or Krebs cycle, the respiratory chain, and ATP synthase (the latter two are also known as the oxidative phosphorylation (OXPHOS) system; Figure 2 and 3). First, glucose is oxidised to pyruvate through glycolysis, which generates 2 ATPs. Then, pyruvate is actively transported into the mitochondrial matrix, where it is converted to acetyl-CoA via the pyruvate dehydrogenase complex (PDHC). Acetyl-CoA is substrate for the TCA cycle, which forms energy-rich metabolites such as flavin adenine dinucleotide, reduced form

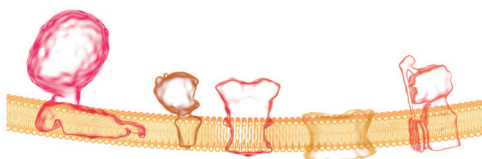


(FADH), nicotinamide adenine dinucleotide, reduced form (NADH) and ATP.<sup>8</sup> In the mitochondrial matrix, the electrons from the TCA cycle are transmitted through NADH to Complex I and subsequently to Coenzyme Q (CoQ<sub>10</sub>). Some electrons are transferred also via FADH and Complex II to CoQ<sub>10</sub>. CoQ<sub>10</sub> transports the electrons to Complex III and via cytochrome c they move to Complex IV where they are ultimately transferred to molecular oxygen to form water. The energy released by the electron transport is used, by Complex I, III and IV, to pump protons into the inter-membrane space, building up a proton gradient across the inner mitochondrial membrane. Complex V or F<sub>0</sub>-F<sub>1</sub> ATP synthase uses this proton gradient for phosphorylation of ADP to ATP.<sup>11</sup> Part of the ATP is used by mitochondria themselves while the majority is transported into the cytosol. Transport of ATP to the cytosol is executed by the ADP/ATP transporter (or adenine nucleotide transporter, ANT).<sup>13</sup>



**Figure 2.**

**Energy production.** Glucose is converted to acetyl-CoA by glycolysis and PDHC. Acetyl-CoA is used in the TCA cycle to produce NADH and FADH, which can be used by the electron transport chain to build up a proton gradient across the mitochondrial inner membrane.

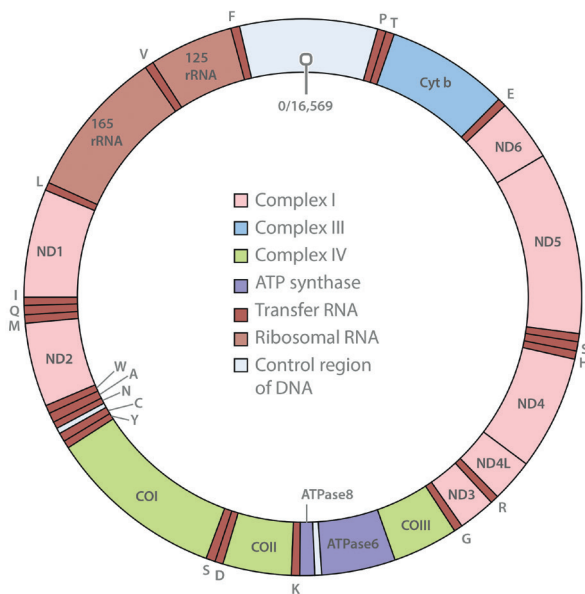


**Figure 3.**

**Artist impression of the oxidative phosphorylation system.** From left to right Complex I, II, III, IV and V.

## Molecular background of mitochondrial disease

The OXPHOS system (Complex I-V; Figure 3) is under dual genetic control, with the exception of Complex II, which is only encoded by nuclear genes. The mitochondrial DNA (mtDNA, Figure 4) encodes for 13 of the 168 structural polypeptides comprising the OXPHOS complexes (subunits of complexes I, III, IV and V), plus 2 ribosomal and 22 transfer RNAs required for the partially autonomous mitochondrial translational apparatus.<sup>14</sup> To date, over 260 pathogenic mutations and 120 large-scale rearrangements have been identified in the mitochondrial genome and more than 170 in nuclear genes.<sup>4</sup> The number of the latter has been steadily increasing since the introduction of whole-exome/whole-genome sequencing.<sup>15</sup>

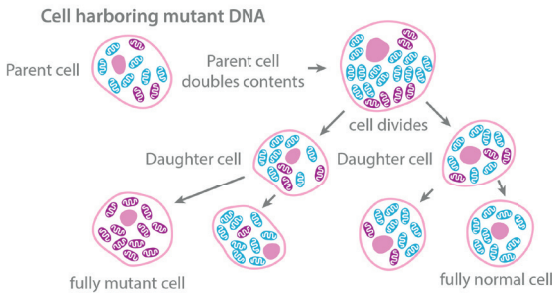


**Figure 4.**  
**Mitochondrial DNA.** The small, circular mtDNA.

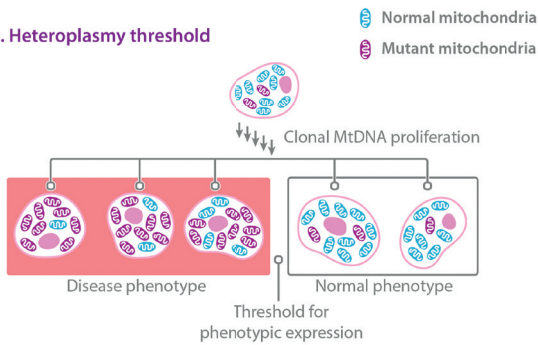
Inheritance of mtDNA mutations has three distinct characteristics: maternal uniparental inheritance, heteroplasmy and the threshold effect (Figure 5).<sup>11</sup> Since mitochondria in the sperm cell are actively destroyed by ubiquitination during fertilisation, mtDNA and therefore mtDNA mutations are exclusively transmitted through the mother, to potentially all her offspring. Since multiple copies of mtDNA are present per cell and mutant mtDNA often co-exists in the cell with normal mtDNA, the mutation may be present in various proportions (heteroplasmy) per mitochondrion and per cell. A high heteroplasmy percentage, as compared to wild-type mtDNA, may lead to biochemical alterations<sup>16</sup>

and signs and symptoms when exceeding a certain threshold. There is no consensus regarding the mutation load causing mitochondrial dysfunction or the possibility of tissue specificity of this threshold, which also differs per mutation. Mitotic segregation, the random distribution of the mutation between daughter cells during mitosis, causes heteroplasmy to vary between organs. This variation between tissues not only complicates the diagnostics,<sup>17,18</sup> but also results in a heterogeneous involvement of organs.

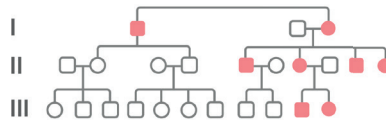
### A. Mitotic segregation



### B. Heteroplasmy threshold



### C. Maternal inheritance



**Figure 5.**

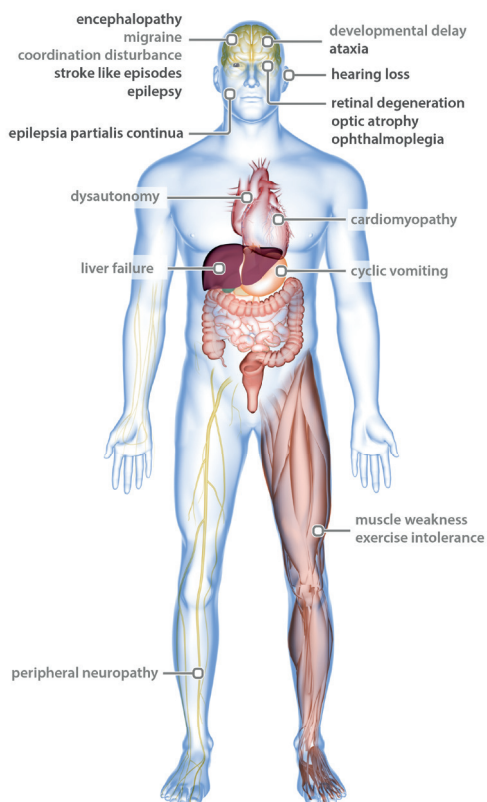
**Hallmarks of mitochondrial inheritance.** When cells proliferate, random distribution of mtDNA across daughter cells (mitotic segregation, A) occurs. For the expression of a phenotype of mitochondrial dysfunction, the heteroplasmy percentage has to exceed a certain threshold (B). Pedigree illustrating maternal inheritance (C). Since mitochondria are inherited via the oocyte, all children of mothers will be at risk of having the mtDNA mutation, whereas the son will not pass the mutation to his children.

Most of the genes involved in mitochondrial function and maintenance have been relocated to the nuclear genome since pro-mitochondria entered the eukaryotic cell approximately 1.5 billions of years ago. Therefore, mutations in nuclear DNA may directly affect components of the oxidative phosphorylation system, by hits in assembly genes, genes encoding proteins involved in mtDNA integrity, replication or mtDNA translation, but also in proteins involved in other functions of the mitochondrion, such as membrane lipid composition, mitochondrial quality control and mitochondrial dynamics.

### How to accomplish the diagnosis of a mitochondrial disorder

The diagnosis of a mitochondrial disease requires the merging of clinical, biochemical and genetic data. Since mitochondria are present in virtually all cells, the clinical diagnostic work-up for a suspected mitochondrial disease includes thorough clinical evaluation, followed by medical history taking which always should include a detailed family history. All commonly affected organs should be screened (physical examination, imaging and electrophysiology) to obtain evidence of the level of multi-organ involvement. The presence of suspicious or highly suspicious symptoms (Figure 6) indicates the clinical probability of a mitochondrial disease. Additional blood, urine and cerebral spinal fluid metabolic investigations may assist in discriminating mitochondrial disorders from other multi-system disorders, although none of the routinely applied biomarkers are 100% specific. Slightly to highly elevated lactate and alanine concentrations in blood and/or CSF are indicative of a mitochondrial disease.<sup>19</sup> Of importance, normal blood lactate concentrations do not rule out a mitochondrial disease.<sup>20</sup> The value of two recently discovered biomarkers, fibroblast growth factor 21 (FGF21) and growth and differentiation factor 15 (GDF15), will be studied in more depth in this thesis.

Next to the clinical and metabolic investigations, the contemporary step in the establishment of a mitochondrial disease is the measurement of the maximal activity of the mitochondrial energy generating system. Mostly, skeletal muscle tissue is used for diagnostic purposes since muscle is a mitochondria-rich tissue and in most patients the skeletal muscles are affected. In patients with predominant liver or heart involvement, a biopsy of the affected tissue might be more informative. Cultured fibroblasts of patients are also analysed in routine diagnostics.<sup>23</sup> To obtain a muscle sample, a needle biopsy of the vastus lateralis dextra muscle is performed in adults while in children preferably a surgically removed muscle biopsy is performed under general anaesthesia.



**Figure 6.**

**Suspicious and highly suspicious symptoms of mitochondrial disease.** In light grey the suspicious symptoms, in dark grey the highly suspicious symptoms.

An alternative to this invasive procedure is the percutaneous conchotome muscle biopsy of the musculus tibialis anterior, which can be performed under local anaesthesia. Geographical breakdown determines whether the sample can be investigated immediately, often termed 'fresh', or should be shipped to a well-equipped mitochondrial centre in frozen condition. Because the reliability of the measurements is higher in fresh muscle tissue and the activity of the whole energy generating system can be measured only in fresh muscle cells, it is recommended to analyse a fresh biopsy.<sup>24,25</sup> Mitochondrial function can be measured applying various strategies; for example, radioactivity based assays to measure the CO<sub>2</sub> production in the presence of different substrates, substrate based-oxygen consumption assays and spectrophotometric analysis of ATP and enzymes and enzyme complexes.<sup>23,24</sup> Following an enzyme deficiency diagnosis, genetic analysis to confirm the biochemical findings is warranted.

In case of a clear mitochondrial syndrome or maternal relatives with a known mtDNA mutation, primary sequencing of the gene suspected to be mutated is recommended. In some conditions, like the m.3243A>G mutation, urine DNA analysis is more sensitive compared to blood analysis.<sup>26,27</sup> The order of the initial investigations (biochemical versus genetic) is currently changing with more rapid and cheaper genetic diagnostics. The biochemical verification of observed genetic alterations, however, is expected to become even more important to confirm the pathogenicity of a (new) mitochondrial disease mutation or mutations in genes of unknown function.

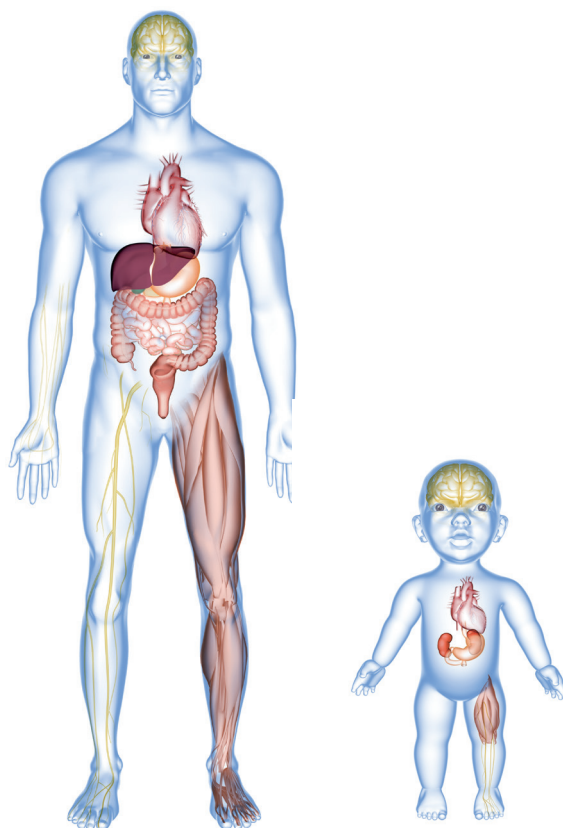
### The phenotype of mitochondrial disorders

Mitochondrial disorders are often said to present 'at any age, with any symptom of any organ and in any mode of inheritance'.<sup>28</sup> In general, mitochondrial diseases are progressive diseases and a substantial number of children with mitochondrial disease dies before puberty. Children presenting at early age generally have the worst prognosis: of those presenting before the age of 6 months, about one third survives beyond their second birthday.<sup>29,30</sup> Children with cardiac involvement, about 17 - 58% in of all paediatric cases,<sup>31-33</sup> have highest morbidity with about 50% of children dying before the age of 12 years.<sup>31</sup> The rate of the decline is unpredictable, even in patients with the same genetic background. There is no clear correlation between the clinical and the biochemical or genetic phenotype.<sup>31,34</sup>

Mitochondrial disorders can be classified based on the molecular defect, the biochemical deficiency or the clinical signs and symptoms. The problem with classification based on molecular defect, is that some defects are very rare and that two patients with the same genetic defect, even those within the same family, may have a completely different phenotype and disease course.<sup>34</sup> Classification based on biochemical deficiencies creates larger groups compared to the stratification based on molecular defects, but the lack of correlation to the clinical phenotype remains. Patients can also be classified based on well-defined mitochondrial syndromes. However, classification based on syndromes also creates small groups with a wide variety in underlying disease causes.<sup>35</sup>

In daily practice, most affected children don't present with a full-blown syndrome, but mostly with a non-specific multi-system phenotype involving tissues requiring most energy, such as the brain, retina, heart, kidney, and skeletal muscle.<sup>4</sup> Roughly, two main clinical phenotypes can be observed: mitochondrial encephalopathy (with mainly central nervous system involvement) and mitochondrial myopathy (mainly muscle involvement). However in practice, most children have a combination of both phenotypes, either or not in combination with other affected organs or tissues (Figure 7). With the discovery of mutations in 'non-classical' mitochondrial pathways, the phenotype associated with mitochondrial disease is rapidly expanding to cover dysmorphias, haematological and endocrine abnormalities, and many more.<sup>15</sup>

In some, but certainly not all patients, a specific mitochondrial syndrome can be observed. Sometimes, the clinical syndrome may give an indication on the genetic cause of the disease, as may be the case in patients with the m.3243A>G mutation associated with for example mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS) or maternally inherited diabetes mellitus and deafness (MIDD).



**Figure 7.**

**Multi-system involvement of mitochondrial disorders.** In patients with mitochondrial disorders, the most commonly affected organs include (top down): brain, eye, ear, heart, liver, gastrointestinal system, nerves and skeletal muscle.

## Mitochondrial myopathy

Mitochondrial myopathy is a heterogeneous entity with variable severity, aetiology and prognosis. Most patients present in early childhood with hypotonia and motor retardation. At later age, symptoms of proximal myopathy leading to a positive Gower's sign, exercise intolerance, ptosis and external ophthalmoplegia may become apparent.

Despite fatigue and sometimes mild concentration problems, intelligence is typically normal. Some patients may be severely limited in their daily activities and fully dependent on wheelchairs and help from caretakers, while others manage to live a reasonably normal life despite their reduced physical capacities.<sup>29</sup> Some patients develop multi-system disease symptoms later in life, including cardiomyopathy, optic atrophy, or sensorineural deafness.

## Mitochondrial encephalopathy

Patients with mitochondrial encephalopathy also present with a very heterogeneous phenotype, ranging from early fatal epilepsy syndromes to migraine and dementia at later age. Children typically present with psychomotor retardation and failure to thrive, either or not in combination with other neurological features such as hypotonia, spasticity and epilepsy.<sup>30</sup> Depression, behaviour problems, and psychosis may complicate the course of the disease.<sup>36</sup> Brain magnetic resonance imaging (MRI) may show specific or non-specific abnormalities,<sup>37</sup> but most patients only have functional defects. Typically, patients are severely mentally retarded, attend a special school and remain dependent on caretakers.<sup>29</sup>

## Leigh syndrome

The prototype of mitochondrial encephalomyopathy is Leigh syndrome (Figure 8). The diagnosis of 'Leigh syndrome' is made on a distinctive MRI picture with a symmetrical distribution of hyperintensities in the basal ganglia, either or not in combination with brainstem, thalamus, diencephalon, cerebellum and spinal cord lesions.<sup>38</sup> Patients usually present within the first months or years of life with psychomotor retardation in combination with signs of brainstem or extrapyramidal dysfunction and lactic acidemia.<sup>38,39</sup> Symptoms include neurological symptoms (muscle weakness, hypotonia, dystonia, respiratory impairment, lethargy, ophthalmoparesis, optic atrophy, ataxia, dysphagia, cranial nerve palsies, and polyneuropathy) sometimes in combination with cardiomyopathy, diarrhoea, vomiting, anaemia or renal problems.<sup>38</sup> Death occurs usually before adulthood as a consequence of respiratory failure caused by on-going brainstem dysfunction either or not in combination with increasing muscle dysfunction.<sup>38</sup> From a functional perspective, these children are mostly wheelchair-bound and severely dependent on their caretakers. For examples, see Box 1.

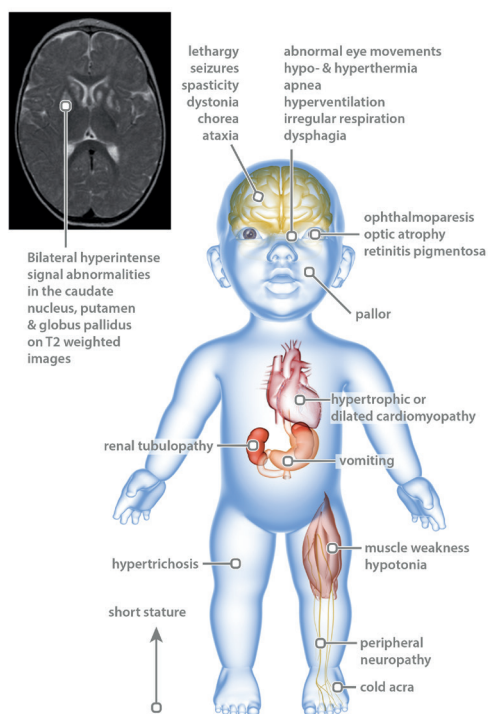
### Box 1: Description of three cases with Leigh syndrome

Rick is a 12-year-old boy suffering from Leigh syndrome due a homozygous mutation in *MTFMT*. His developmental age is approximately 2 years. He is sociable and able to communicate in short sentences. He is able to follow simple and single commands but his cooperation, attention span and understanding are limited. His neurological examination shows strabismus, decreased facial expression, proximal muscle weakness and mild spasticity in his legs. He has nocturnal ventilation and is not able to go to school due to tiredness.



Jack is a 12-year-old boy with Leigh syndrome due to a homozygous mutation in *NDUFS7*. He is able to communicate via eye movements but not through speech. Except for some involuntary movements of the face and his left arm, he is not able move. He has severe contractures of his joints and a mixed picture of atonia and spasticity. He is fed through a gastrostomy and has to be ventilated 24 hours a day because of central hypoventilation.

Peter was 2 years old and carries a heterozygous mutation in *NDUFS7*. During an infectious episode at the age of 15 months, he was diagnosed with Leigh syndrome, based on a severe hypotonia and a characteristic MRI picture. After this period, his condition improved and he developed normally. The only abnormal sign on physical examination was an intention tremor. At the age of 28 months, his condition deteriorated acutely and he was admitted because of hypoventilation. A week after the admission, he died from a central apnoea.



**Figure 8.**

**Leigh syndrome.** Cartoon depicting symptoms that can be observed in patients with Leigh syndrome.

## The m.3243A>G mutation

Another mitochondrial disease entity illustrating the heterogeneity of mitochondrial disorders, are those caused by the 3243A>G nucleotide change in the mitochondrial DNA transfer RNA of Leucine (Figure 9). This point-mutation, classically known as the MELAS mutation, actually has a much wider spectrum of clinical phenotypes. Whereas high heteroplasmy levels are indeed associated with mitochondrial encephalopathy, lactic acidosis and stroke-like episodes (MELAS) syndrome, low to moderate heteroplasmy levels are associated with a huge variety of other phenotypes.<sup>40,41</sup> Examples hereof include mitochondrial inherited diabetes and deafness (MIDD), cardiomyopathy, macular abnormalities, and myopathy. Where some patients may only become aware of the cause of their mild complaints through the diagnosis of a maternal family member, other patients present in early infancy and remain dependent on their caretakers life-long. For examples, see Box 2.

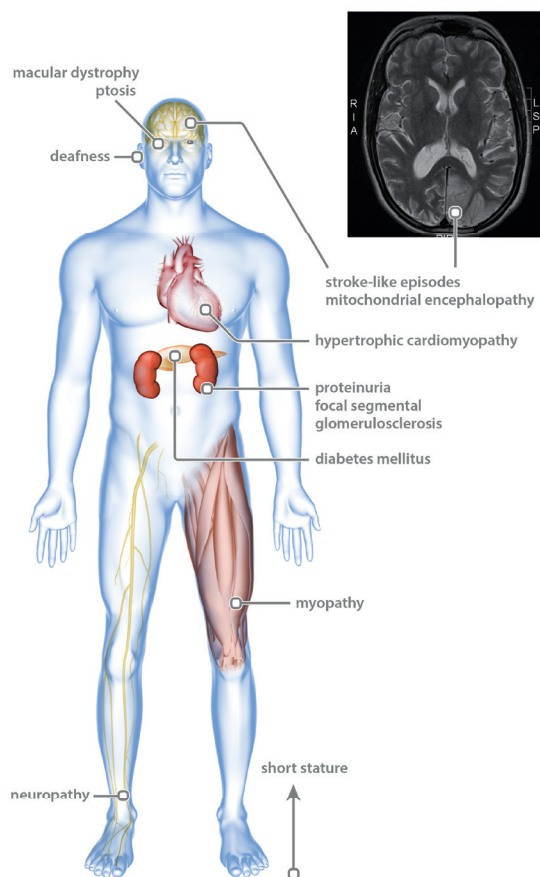
### **Box 2: Description of three cases with a m.3243A>G mutation.**

Ralph is a 32-year-old man who started suffering from progressive hearing loss at the age of 15 years. He had a history of slow motor and cognitive development. Six years after his first symptoms, he suffered from his first stroke-like episode with confusion, headache, and aphasia, followed by two stroke-like episodes at the age of 31 and 34 years of age. After the first stroke-like-episode, his cognitive abilities declined and he suffered from epilepsy and migraine. At the age of 28 years, a severe hypertrophic cardiomyopathy with cardiac failure was detected, which only slowly improved after medication. A few years later, at the age of 30 years, he was admitted because of a severe diabetic ketoacidosis as a consequence of a diabetes de novo. The m.3243A>G mutation (>90% in muscle and 60% in leucocytes) was found as an explanation for his complaints. At this moment, Ralph is severely impaired in his activities of daily living and suffers from psychotic episodes.

Rose, the 34-year-old sister of Ralph, was found to carry the m.3243A>G mutation (heteroplasmy percentage 40%) when her brother was diagnosed with MELAS syndrome. She suffers from migraine headaches, insulindependent diabetes mellitus, moderate deafness, and constipation. She has a very mild hypertrophic cardiomyopathy, but has no cardiac complaints. At physical examination, muscle atrophy and decreased muscle power are found. Her retina shows some hyperpigmentation. She contemplates prenatal genetic diagnostics for having a child on her own.

The 60-years-old mother of Rose and Ralph has a heteroplasmy level of 20%. Her hearing is normal, and she has no diabetes mellitus, cardiomyopathy or myopathic complaints.

Tom is the 9-year-old nephew of Ralph. Despite that he is not very good at school gymnastics and it took him 3.5 years to get his first swimming certificate, he has no complaints so far.



**Figure 9.**

**Symptoms that can be observed in patients with the m.3243A>G mutation.** A) MELAS syndrome; B) MIDD. Note the overlap between MIDD and MELAS syndrome.

## Mitochondrial Medicine: the long Road to Treatment

The term mitochondrion was proposed by Benda in 1898 to describe a group of subcellular organelles with faint, threadlike granules.<sup>42</sup> The widely accepted endosymbiont theory states that these small organelles resulted from the primordial ingestion of the energy-efficient and oxygen-detoxifying prokaryotic bacteria-like protomitochondria by prokaryotic cells.<sup>43</sup> The importance of these organelles in metabolic function was only defined in the early 20<sup>th</sup> century through a series of experiments describing the role of these organelles in oxidative phosphorylation.<sup>44</sup>

The first described patient with a mitochondrial disease was a young Swedish woman with severe hypermetabolism with electron microscopically abnormal mitochondria and loose coupling of oxidation and phosphorylation on examination of the muscle, reported by Ernster and Luft in 1962.<sup>45</sup> This paper introduced the concept of mitochondrial medicine and numerous biochemical and histochemical assays allowing detection of mitochondrial function were developed and used to diagnose patients with mitochondrial disorders. Soon, the heterogeneity and multi-system nature of respiratory chain disorders became clear and Shapira introduced the term 'mitochondrial encephalomyopathies' to indicate disorders in which the oxidative phosphorylation was hampered.<sup>46</sup> Several clinical syndromes were defined, such as MELAS and MERRF (myoclonus epilepsy with ragged red fibres) syndrome.

Although mitochondrial DNA (mtDNA) was first described in 1967, it was not until 1988 that the first disease-causing mutation in this small molecule was discovered. The group of Harding and colleagues found the first large-scale single deletions in mtDNA in patients with mitochondrial myopathies,<sup>47</sup> quickly followed by the discovery of a single gene mutation in mtDNA causing LHON.<sup>48</sup> That first discovery was soon followed by many others and the mitochondrial disease 'molecular era' had come.<sup>49</sup> In the next decade, when approximately 100 mutations in mtDNA were listed, the first mutation in nuclear DNA was described in the small and nuclear encoded second complex of the respiratory chain, followed by the first mutation in Complex I by the Nijmegen group.<sup>50,51</sup>

Increased knowledge about the structure, function and pathology of the OXPHOS system has led to some obvious intervention strategies. In 1989, the first patients with a primary coenzyme Q<sub>10</sub> (CoQ<sub>10</sub>) biosynthesis defect was described. The first description reported some improvement in fatigue and alertness on CoQ<sub>10</sub> supplementation (150 mg)<sup>52</sup> whereas later studies report full recovery (up to 300 mg).<sup>31</sup> Since then, only ~75 cases have been described in literature, despite the systematic screening for these defects in muscle biopsies.<sup>53</sup> More recently, allogeneic haematopoietic stem cell transplantation and thereby reintroduction of the thymidine phosphorylase enzyme, was shown to be able to lower the toxic and disease causing thymidine and deoxyuridine

levels in blood in patients with mitochondrial neuro-gastro-intestinal encephalopathy (MNGIE).<sup>54,55</sup> The promising results of successful allogeneic haematopoietic stem cell transplantation are however hampered by the high mortality rate of the procedure. Finally, stroke-like episodes in patients with MELAS syndrome were suggested to be caused by endothelial dysfunction.<sup>56</sup> Since endothelial function is regulated by nitric oxide, a precursor of nitric oxide, L-arginine, was infused. This resulted in improvement of headache and nausea during stroke-like episodes, prevention of stroke-like episodes when used prophylactically and in an increased total body maximal aerobic capacity.<sup>55-59</sup> Unfortunately, all of these observations are non-blinded and performed in just a handful of patients.

The above mentioned strategies only apply to a small proportion of all subjects with mitochondrial disease. Various new strategies, which may improve the lives of many more mitochondrial disease patients in the future, are under investigation. These strategies include: i) prevention of transmission of gene defects to offspring; ii) replacement or repair of defective genes (gene therapy); iii) altering the balance between wild-type and mutated mtDNA; iv) controlled regulation of specific transcriptional regulators; v) inhibition of autophagy; and vi) metabolic manipulation.<sup>60</sup>

Preventing transmission of known disease-causing mutations and/or known respiratory chain enzyme defects, by prenatal biochemical/genetic or pre-implantation genetic diagnostics or oocyte donation, is part of daily care.<sup>61,62</sup> Mitochondrial donation, allowing women with mtDNA mutations to have children with their own nuclear DNA but with healthy mitochondria, has been approved by the House of Lords of the United Kingdom in 2015, is still in its infancy.

Gene therapy, covering both the second and the third strategy, aims at i) targeted re-expression of the mutated gene to correct single mutations in either nuclear or mtDNA; ii) at manipulation of the heteroplasmy percentage using zinc fingers or transcription Activator-Like Effector Nucleases; or iii) at overexpression of proteins involved in the stability of mt-tRNA in case of mutations in genes encoding mt-tRNAs. Clinical trials testing the effectiveness of gene therapy, mostly in (non-mitochondrial) ocular disorders, are currently running.

The expression and amount of mitochondrial OXPHOS proteins (approach iv) can be influenced by e.g. transcription regulation or endurance training.<sup>63</sup> For example, pathways controlling mitochondrial biogenesis can be targeted via peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ ) coactivators (PGC family), which activates several transcription factors regulating the expression of OXPHOS genes. Both NAD-dependent protein deacetylase sirtuin 1 (SIRT1) and 5'adenosine monophosphate-activated protein kinase (AMPK), enzymes regulating the activity of PGC1 $\alpha$ , the most well characterised enzyme

of the PGC family, are drugable enzymes. In a mouse model for mitochondrial disease, the upregulation of mitochondrial biogenesis by these activators showed improvement of both biochemical and clinical parameters.<sup>64,65</sup> However, because of a lack of specificity of the consequences of transcription regulation, it is highly likely that more side effects apart from the already recognised hepatic steatosis will occur.<sup>66</sup> Healthy subjects receiving the AMPK activator R118 showed altered mental status, hemodynamic instability and increased blood lactate concentrations, which was thought to be caused by mitochondrial dysfunction instead of increased mitochondrial function.<sup>67</sup>

Activation of autophagy (approach v), the recycling of degraded cellular organelles to warrant their quality, can be inhibited by rapamycin (sirolimus). The administration of rapamycin to a mouse model for mitochondrial disease both delayed disease progression and fatal outcome, and corrected several of the biochemical abnormalities observed in this mouse model.<sup>68</sup> Also for this strategy, the long-term effects on organelle quality are unknown and more investigations are indicated to elucidate the long term effect and side effects of rapamycin.

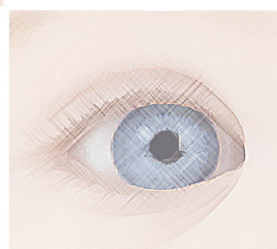
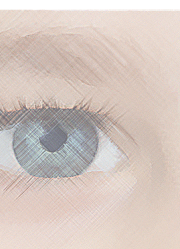
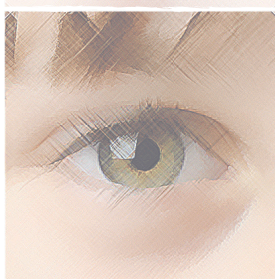
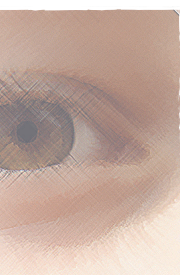
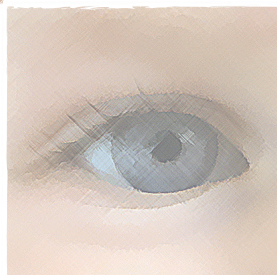
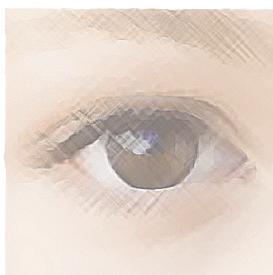
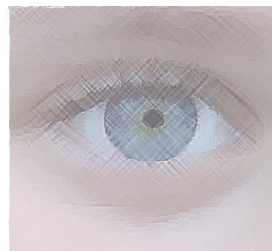
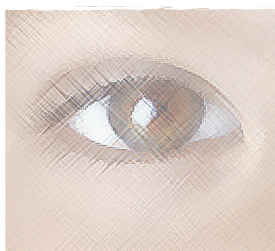
Metabolic manipulation by the ketogenic diet aims at stimulating mitochondrial  $\beta$ -oxidation and providing ketones as an alternative substrate for brain, heart and skeletal muscle. This approach has shown variable effects in mice<sup>69,70</sup> but was poorly tolerated by patients with mitochondrial myopathy, who uniformly showed severe muscle pain with elevated creatine kinase followed by progressive increase of serum myoglobin values.<sup>71</sup>

The sixth approach also aims at counteracting specific derangements in cellular homeostasis caused by mitochondrial dysfunction.<sup>4,60,72</sup> Well-characterised alterations include increased reactive oxygen species production, altered redox-metabolism, alterations in the mitochondrial network architecture, diminished mitochondrial membrane potential and altered cellular mitochondrial calcium handling.

The most frequently studied consequence of mitochondrial dysfunction is an increased production of reactive oxygen species or radicals (ROS). ROS are a physiological by-product of metabolism and play an important physiological role in protein expression and regulation via various signalling pathways.<sup>73</sup> If ROS production becomes higher than can be counterbalanced by the cell's own anti-oxidant system,<sup>74,75</sup> damage to proteins, lipids and DNA may occur, which is thought to have harmful effects on cellular and mitochondrial (ultra)structure, activity and matrix protein diffusion.<sup>3,74-77</sup> To what extent ROS are involved in the pathophysiology of mitochondrial disease is still under debate.<sup>74,75</sup> Despite the on-going controversies of ROS' role in mitochondrial disease pathology, much effort is devoted to studying the effects of lowering (scavenging) ROS in mitochondrial disease models and patients. The clinically most well-studied scavenger

is CoQ<sub>10</sub>, which acts as an electron transporter in the mitochondrial respiratory chain where it transports electrons from Complex I and II to Complex III.<sup>78</sup> The results of CoQ<sub>10</sub> supplementation in mitochondrial disorders other than CoQ<sub>10</sub> biosynthesis defects are less convincing.<sup>79</sup> A synthetic CoQ derivative, Idebenone (hydroxydecyl benzoquinone), is well studied in Leber Hereditary Optic Neuropathy (LHON),<sup>80,81</sup> a mitochondrial disease that causes blindness in young adults as well as in Friedreich's ataxia, a progressive, multi-system neurodegenerative disorder.<sup>82,83</sup> On both diseases, no significant effect of idebenone was observed. EPI-743, another CoQ<sub>10</sub> analogue combined with vitamin E, was shown to be safe in several open-label safety studies.<sup>73,74,84</sup> These safety studies also report reversion of vision loss in six patients with LHON<sup>85,86</sup> and improvement of gross motor function and general mitochondrial disease severity in 22 children with mitochondrial disease.<sup>84</sup> The same company is currently testing another drug candidate, EPI-589, in a phase 1B.<sup>87</sup> Bendavia, a drug targeted at cardiolipin within the inner mitochondrial membrane, also aims at reducing ROS production. The topical solution containing Bendavia, Ocuvia, was designed to treat ocular disorders such as LHON. Safety studies for both drugs are currently ongoing.<sup>87</sup> Khondrion, located in Nijmegen, has developed KH176, a drug influencing ROS-redox metabolism, which showed a favourable safety profile in recent phase 1 study.<sup>87</sup>

Since no cure for most mitochondrial disorders is available yet, present day management is focussed on secondary prevention, symptomatic treatment, rehabilitation, and psychological support. The management of a patient with a mitochondrial disease includes life-long advices regarding nutrition, medication, exercise, and recurrence risk as well as screening for common treatable complications of mitochondrial disease, such as cardiomyopathy, renal failure, and cardiac conduction defects. All of these preventive interventions still rely on empirical data and expert opinions.





# Problem definition

Patients, their relatives and physicians are desperate to find any treatment that helps to cure or alleviate mitochondrial diseases. Until 1990, when the first randomised, placebo controlled clinical trial in mitochondrial diseases was performed,<sup>88</sup> the 'treatment' of patients was mostly based on rationally selected vitamins of which anecdotal evidence of effectiveness was available. These open-label studies typically involved a heterogeneous group of just a handful of patients reporting improved quality of life with the use of these vitamins. Since there were no obvious side effects, the need for evidence regarding their efficacy was never high on the agenda. In 2006, the first critical systematic review of the literature stated that most of the studies in patients with mitochondrial disease had inadequate trial design or a non-conclusive outcome.<sup>89</sup> A few years later the number of trials had doubled, but the percentage of well designed and reported randomised clinical trials remained stable and extremely low (12 out of 1,335 studies).<sup>5</sup> This was the starting point for a more critical attitude towards open-label studies in small groups and for the discussion on how to improve the quality of future studies.

Designing a clinical study to prove treatment effects in patients, and especially children, with mitochondrial disorders is challenging for many reasons (Box 3). Mitochondrial disorders are rare, heterogeneous diseases involving multiple organs and the clinical disease course is unpredictable, variable and often oscillating. Although the variability in affected organs and degree of handicap (illustrated by the cases in Box 1 and 2) may warrant grouping of patients into homogeneous clinical cohorts, both the external validation of the results of the trial and the potential to detect clinical benefit for a small group of patients with specific biochemical or genetic abnormalities are hampered when setting very strict inclusion criteria. Lastly, the unpredictable disease course requires such large cohorts that patient recruitment will be impossible despite the participation of multiple centres.

### **Box 3: Hurdles in proving treatment effects in patients with mitochondrial disease**

Rare disorders

Clinically, biochemically and genetically heterogeneous disorders

Clinical heterogeneity within one genetic entity, both in mtDNA and nDNA mutations

Unpredictable, variable and sometimes oscillating natural disease course

Multi-system disease with variable involvement of organs

Patients are often very disabled and unable to perform functional tests

Abilities of patients vary widely

#### **Child-specific aspects**

Children are developing and growing

Outcome measures for children mostly have a limited age range

Paediatric patients are often unable to communicate

Children have a shorter attention span and are less likely to follow the instructions completely

The cooperation required for some tests is highly dependent on the level of enthusiasm for the test and the mental abilities of the child

The use of robust end points relevant to and feasible for a large proportion of patients with mitochondrial disease may assist in widening the inclusion criteria of the trial and therefore increase its feasibility and the generalizability of the results. The end points that have been used so far have rarely been validated and do not always measure symptoms and complaints experienced by or relevant to the patient. Therefore, we sought to improve this aspect of clinical trial design by looking for reproducible, sensitive and responsive clinical biomarkers indicative of disease severity.

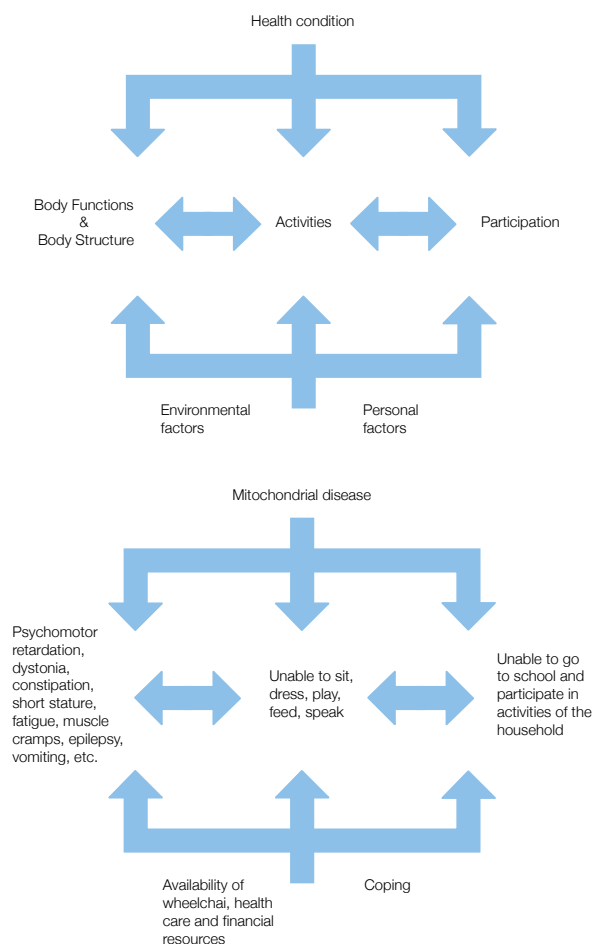
## Outcome measure selection for trials in patients with mitochondrial disease

In medicine, outcome measures are defined as instruments providing information about certain disease aspects, such as disease severity, response to therapy or remission of the disease.<sup>90</sup> It is said that 'Clinical trials are only as credible as their end points',<sup>91</sup> since the success of the trial is based on the change in the primary end point. At this moment, the selection of outcome measures for clinical trials evaluating interventions with new drugs in patients with mitochondrial disorders is inefficient and not systematically addressed. As only those trials in which end points are assessed with valid instruments provide meaningful data, the selection of a robust instrument for the targeted population should have the highest priority. Since the mitochondrial disease field has paid little attention to the selection of outcome measures, the selection and use of robust instruments will - as we expect - lead to an enormous increase in the quality of trials and significance of data. Here, I will elaborate on the strategy to identify and select instruments measuring clinically relevant aspects of health in children with mitochondrial disorders.

## Identification of available instruments

Surely, the improvement in health and disability after the intended treatment should be experienced by and relevant to the patient. This is endorsed by regulatory authorities such as the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA), preferring the inclusion of clinically meaningful outcomes in clinical trials. How can we identify these relevant aspects of health to patients with mitochondrial disease?

According to the World Health Organisation (WHO), health is a multi-dimensional construct: 'a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'. Forty-six years after the rather non-specific but holistic definition of health, the WHO developed a framework for thinking about health in chronic conditions, taking into account the bio psychosocial aspect of human functioning.<sup>92</sup> This framework, the International Classification of Functioning, Disability and Health (ICF, Figure 10) with a separate version for children and youth (ICF-CY), comprises four dimensions: i) body functions; ii) body structure; iii) activities and participation; and iv) environmental factors.



**Figure 10.**

**The International Classification of Functioning, Disability and Health (ICF) framework.** The structure of the ICF (left), filled out as an example for the symptoms, limitations and disabilities that can be observed in patients with mitochondrial disease (right). Note the interplay between the health condition on physical function, activity and participation on one side, and the influence of environmental and personal factors on the other.

In the ICF-model, the severity of the experienced disability is seen as a dynamic, interactive process among the interconnected components of functioning in contextual factors.<sup>93</sup> For example, an increase in muscle strength may not correlate with the daily functioning<sup>94</sup> and may be irrelevant to a patient not able to adequately coordinate movements. To systematically identify the complaints and limitation relevant to all, or at least a large majority of patients with mitochondrial disorders, the framework provided by the ICF may be of major assistance.

## Selection of instruments

When the relevant aspects of the disease are known, instruments measuring the functional and structural consequences of mitochondrial disease can be identified, as well as functional tests measuring the consequences of the disease on activities and participation. The ICF-framework fails to identify two groups of instruments, namely those measuring quality of life and biomarkers. The latter should be indicative of disease severity and should change with disease progression or regression to be applicable as an outcome measure in clinical trials.

Subsequent selection of the most adequate instrument should be based on the internal properties of the instrument in the intended population. Validation studies seek to explore these internal properties of a certain instrument, including feasibility, reliability, validity and responsiveness (Table 1 and Figure 11).<sup>95-99</sup> The reliability (precision) of an instrument is the extent to which scores for patients who have not changed are the same for repeated measurements under several conditions, using different sets of items from the same multi-item measurement instrument (internal consistency), over time (test-retest), by different raters on the same occasion (inter-rater) and by the same rater on different occasions (intra-rater).<sup>99</sup> The test-retest reliability of an outcome measure is determined by the variability in the subject, in the observer, and in the measurement itself. In practice, it is difficult to separate these sources of error. In inter- and intra-rater experiments, the variation is mainly related to the observer and the measurement itself. The validity (accuracy) of a measurement instrument answers the question whether an instrument measures the concept it is supposed to measure.<sup>99</sup> The validation process consists of a series of studies. First, feasibility, face validity ('looks like' measuring what it is supposed to measure) and content validity (measures useful items) are assessed for the targeted patient group. If the instrument passes through the first validity tests, criterion validity is assessed by correlating the measure with a gold standard. A high correlation means that the instrument demonstrates the real value of the construct measured. Since for many disorders no gold standard is available, correlation with other well-validated measures is often performed (construct validity). Construct validity can be broken down into two subcategories: convergent validity, agreement with previous instruments rating the same aspect of disease and divergent validity, whether this instrument is not correlated to instrument with which there is a lack of relationship. The responsiveness of an instrument measures the ability to detect changes over time. Usually, the relationship between the evaluated instrument and other end points is evaluated in a treatment setting. Since we don't have effective treatment yet, responsiveness will be difficult to determine. The assessment of an instrument should approximately follow this sequence, since unfeasible and unreliable instruments need no assessment of validity.

Table 1. Internal properties of an instrument: explanation.

<b>Reliability = the likelihood that a measurement reflects the true value of the measurement instead of random error</b>	
Interna consistency	Consistency of subitems in measuring the same concept
Intra-rater reliability	Repeatability by the same rater on different occasions
Inter-rater reliability	Reproducibility between observers on the same occasion
Test-retest reliability	Reproducibility within the same subject over time
<b>Validity = generalizability of interpretation, measures a stable concept</b>	
Face validity	Looks like measuring the right construct
Content validity	Relevance and comprehensiveness of items in measuring the construct
Criterion validityw	Reflection of value of the construct (i.e. gold standard)
Construct validity	Reflection of the dimensionality of the construct measured
Cross-cultural validity	Performance of items of a translated or culturally adapted instrument
<b>Responsiveness = the ability to detect change over time in the construct measured</b>	
Criterion responsiveness	Detection of change of the construct (i.e. gold standard)
Construct responsiveness	Detection of the change in a derivative of the construct
<b>Responsiveness = the ability to detect change over time in the construct measured</b>	
Discriminative ability	Discriminative ability between healthy and diseased persons
Floor effect	Discriminative ability in low scores
Ceiling effect	Discriminative ability in high scores
Smallest detectable change	Smallest change in score that can be measured in an individual
Minimum important change	Smallest change in score patients perceive as important

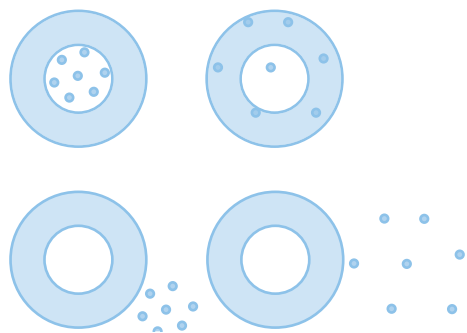


Figure 11. Illustration of reliability (precision) and validity (accuracy). A) Both accuracy and precision; B) accuracy only; C) precision only; D) Neither accuracy nor precision.

It is highly unlikely that the perfect instrument, which is feasible, highly relevant and 100% precise and accurate in all children with any mitochondrial disorder exists. However, the systematic appraisal of available instruments and the subsequent selection of most robust instruments based on data instead of gut feeling will certainly increase the quality of the information obtained by future clinical trials.

## References

1. Chinnery PF, Turnbull DM. Epidemiology and treatment of mitochondrial disorders. *Am J Med Genet* 2001;106:94-101.
2. Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders-past, present and future. *Biochimica et biophysica acta* 2004;1659:115-20.
3. Smeitink JA, Zeviani M, Turnbull DM, Jacobs HT. Mitochondrial medicine: a metabolic perspective on the pathology of oxidative phosphorylation disorders. *Cell metabolism* 2006;3:9-13.
4. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med* 2012;366:1132-41.
5. Pfeffer G, Majamaa K, Turnbull DM, Thorburn D, Chinnery PF. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2012;4:CD004426.
6. Henze K, Martin W. Evolutionary biology: essence of mitochondria. *Nature* 2003;426:127-8.
7. Koopman WJ, Visch HJ, Verkaart S, van den Heuvel LW, Smeitink JA, Willems PH. Mitochondrial network complexity and pathological decrease in complex I activity are tightly correlated in isolated human complex I deficiency. *American journal of physiology* 2005;289:C881-90.
8. Alberts B, Bray T, Johnson A, *et al.* *Essential Cell Biology: An introduction to the Molecular Biology of the Cell*. New York: Garland Publishing; 1998.
9. Mannella CA. Structure and dynamics of the mitochondrial inner membrane cristae. *Biochimica et biophysica acta* 2006;1763:542-8.
10. Harner M, Korner C, Walther D, *et al.* The mitochondrial contact site complex, a determinant of mitochondrial architecture. *Embo J* 2011;30:4356-70.
11. Smeitink J, van den Heuvel L, DiMauro S. The genetics and pathology of oxidative phosphorylation. *Nature reviews* 2001;2:342-52.
12. McBride HM, Neuspiel M, Wasiak S. Mitochondria: more than just a powerhouse. *Current biology* : CB 2006;16:R551-60.
13. Talbot DA, Duchamp C, Rey B, *et al.* Uncoupling protein and ATP/ADP carrier increase mitochondrial proton conductance after cold adaptation of king penguins. *J Physiol* 2004;558:123-35.
14. Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nature reviews* 2005;6:389-402.
15. Fung CW, Koene S, Willems PHGM, Koopman WJH, Smeitink J. Mitochondrial disorders: Emerging biogenesis pathways, novel clinical phenotypes and treatment strategies. *EMBO molecular medicine* 2015.
16. Janssen AJ, Schuelke M, Smeitink JA, *et al.* Muscle 3243A>G mutation load and capacity of the mitochondrial energy-generating system. *Annals of neurology* 2008;63:473-81.
17. Rahman S, Poulton J, Marchington D, Suomalainen A. Decrease of 3243 A>G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *American journal of human genetics* 2001;68:238-40.
18. Ciafalonì E, Ricci E, Shanske S, *et al.* MELAS: clinical features, biochemistry, and molecular genetics. *Annals of neurology* 1992;31:391-8.
19. Morava E, Hogeveen M, De Vries M, Ruitenbeek W, de Boode WP, Smeitink J. Normal serum alanine concentration differentiates transient neonatal lactic acidemia from an inborn error of energy metabolism. *Biology of the neonate* 2006;90:207-9.



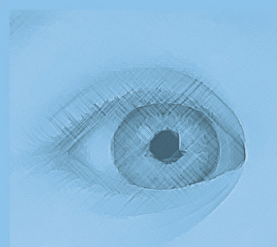
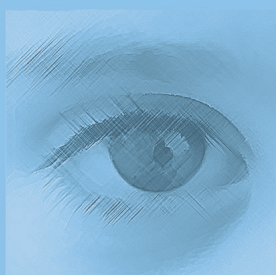
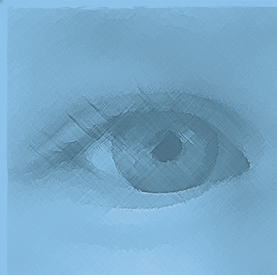
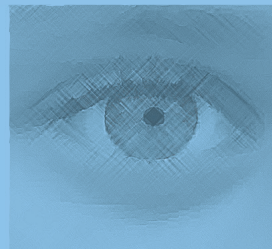
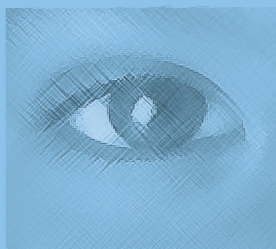
20. Robinson BH. Lactic acidemia and mitochondrial disease. *Molecular genetics and metabolism* 2006;89:3-13.
21. Suomalainen A, Elo JM, Pietilainen KH, *et al.* FGF-21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. *Lancet Neurol* 2011;10:806-18.
22. Kalko SG, Paco S, Jou C, *et al.* Transcriptomic profiling of TK2 deficient human skeletal muscle suggests a role for the p53 signalling pathway and identifies Growth and Differentiation Factor-15 as a potential novel biomarker for mitochondrial myopathies. *BMC genomics* 2014;15:91.
23. Janssen AJ, Tjibbels FJ, Sengers RC, *et al.* Spectrophotometric assay for complex I of the respiratory chain in tissue samples and cultured fibroblasts. *Clinical chemistry* 2007;53:729-34.
24. Janssen AJ, Tjibbels FJ, Sengers RC, *et al.* Measurement of the energy-generating capacity of human muscle mitochondria: diagnostic procedure and application to human pathology. *Clinical chemistry* 2006;52:860-71.
25. Janssen AJ, Smeitink JA, van den Heuvel LP. Some practical aspects of providing a diagnostic service for respiratory chain defects. *Ann Clin Biochem* 2003;40:3-8.
26. de Laat P, Koene S, van den Heuvel LP, Rodenburg RJ, Janssen MC, Smeitink JA. Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A>G mutation. *Journal of inherited metabolic disease* 2012.
27. Whittaker RG, Blackwood JK, Alston CL, *et al.* Urine heteroplasmy is the best predictor of clinical outcome in the m.3243A>G mtDNA mutation. *Neurology* 2009;72:568-9.
28. Munnich A, Rotig A, Chretien D, *et al.* Clinical presentation of mitochondrial disorders in childhood. *Journal of inherited metabolic disease* 1996;19:521-7.
29. Debray FG, Lambert M, Chevalier I, *et al.* Long-term outcome and clinical spectrum of 73 pediatric patients with mitochondrial diseases. *Pediatrics* 2007;119:722-33.
30. Skladal D, Sudmeier C, Konstantopoulou V, *et al.* The clinical spectrum of mitochondrial disease in 75 pediatric patients. *Clin Pediatr (Phila)* 2003;42:703-10.
31. Scaglia F, Towbin JA, Craigen WJ, *et al.* Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics* 2004;114:925-31.
32. Holmgren D, Wahlander H, Eriksson BO, Oldfors A, Holme E, Tulinius M. Cardiomyopathy in children with mitochondrial disease; clinical course and cardiological findings. *Eur Heart J* 2003;24:280-8.
33. Darin N, Oldfors A, Moslemi AR, Holme E, Tulinius M. The incidence of mitochondrial encephalomyopathies in childhood: clinical features and morphological, biochemical, and DNA abnormalities. *Annals of neurology* 2001;49:377-83.
34. Koene S, Rodenburg RJ, van der Knaap MS, *et al.* Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: what we learned from 130 cases. *Journal of inherited metabolic disease* 2012.
35. Leigh D. Subacute necrotizing encephalomyelopathy in an infant. *Journal of neurology, neurosurgery, and psychiatry* 1951;14:216-21.
36. Koene S, Kozicz TL, Rodenburg RJ, *et al.* Major depression in adolescent children consecutively diagnosed with mitochondrial disorder. *Journal of affective disorders* 2009;114:327-32.
37. Valanne L, Ketonen L, Majander A, Suomalainen A, Pihko H. Neuroradiologic findings in children with mitochondrial disorders. *AJNR Am J Neuroradiol* 1998;19:369-77.
38. Finsterer J. Leigh and Leigh-like syndrome in children and adults. *Pediatr Neurol* 2008;39:223-35.

39. Rahman S, Blok RB, Dahl HH, *et al.* Leigh syndrome: clinical features and biochemical and DNA abnormalities. *Annals of neurology* 1996;39:343-51.
40. Sproule DM, Kaufmann P. Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes: basic concepts, clinical phenotype, and therapeutic management of MELAS syndrome. *Annals of the New York Academy of Sciences* 2008;1142:133-58.
41. Laloi-Michelin M, Meas T, Ambonville C, *et al.* The clinical variability of maternally inherited diabetes and deafness is associated with the degree of heteroplasmy in blood leukocytes. *J Clin Endocrinol Metab* 2009;94:3025-30.
42. Benda C. Ueber die Spermatogenese der Vertrebraten und hoherer Evertrebraten, II. Theil: Die Histiogenese der Spermien. *Arch Anat Physiol* 1898;73:393-8.
43. Sagan L. On the origin of mitosing cells. *Journal of theoretical biology* 1967;14:255-74.
44. Schon EA. The Mitochondrial Genome. In: Rosenberg R, ed. *The Molecular and Genetic Basis of Neurologic and Psychiatric Disease* 3rd ed. Philadelphia, PA: Butterworth-Heinemann; 2003:179-88.
45. Luft R, Ikkos D, Palmieri G, Ernster L, Afzelius B. A case of severe hypermetabolism of nonthyroid origin with a defect in the maintenance of mitochondrial respiratory control: a correlated clinical, biochemical, and morphological study. *J Clin Invest* 1962;41:1776-804.
46. Shapira Y, Harel S, Russell A. Mitochondrial encephalomyopathies: a group of neuromuscular disorders with defects in oxidative metabolism. *Isr J Med Sci* 1977;13:161-4.
47. Holt IJ, Cooper JM, Morgan-Hughes JA, Harding AE. Deletions of muscle mitochondrial DNA. *Lancet* 1988;1:1462.
48. Wallace DC, Singh G, Lott MT, *et al.* Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* 1988;242:1427-30.
49. Dimauro S. A history of mitochondrial diseases. *Journal of inherited metabolic disease* 2011.
50. Bourgeron T, Rustin P, Chretien D, *et al.* Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. *Nature genetics* 1995;11:144-9.
51. van den Heuvel L, Ruitenbeek W, Smeets R, *et al.* Demonstration of a new pathogenic mutation in human complex I deficiency: a 5-bp duplication in the nuclear gene encoding the 18-kD (AQDQ) subunit. *American journal of human genetics* 1998;62:262-8.
52. Ogasahara S, Engel AG, Frens D, Mack D. Muscle coenzyme Q deficiency in familial mitochondrial encephalomyopathy. *Proc Natl Acad Sci U S A* 1989;86:2379-82.
53. Desbats MA, Lunardi G, Doimo M, Trevisson E, Salvati L. Genetic bases and clinical manifestations of coenzyme Q (CoQ ) deficiency. *Journal of inherited metabolic disease* 2014;5:5.
54. Halter J, Schupbach WM, Casali C, *et al.* Allogeneic hematopoietic SCT as treatment option for patients with mitochondrial neurogastrointestinal encephalomyopathy (MNGIE): a consensus conference proposal for a standardized approach. *Bone Marrow Transplant* 2011;46:330-7.
55. Hirano M, Garone C, Quinzii CM. CoQ(10) deficiencies and MNGIE: Two treatable mitochondrial disorders. *Biochimica et biophysica acta* 2012;1820:625-31.
56. Koga Y, Povalko N, Nishioka J, Katayama K, Yatsuga S, Matsuishi T. Molecular pathology of MELAS and L-arginine effects. *Biochimica et biophysica acta* 2012;1820:608-14.
57. Koga Y, Akita Y, Nishioka J, *et al.* L-arginine improves the symptoms of stroke-like episodes in MELAS. *Neurology* 2005;64:710-2.
58. Siddiq I, Widjaja E, Tein I. Clinical and radiologic reversal of stroke-like episodes in MELAS with high-dose L-arginine. *Neurology* 2015.

59. Rodan LH, Wells GD, Banks L, Thompson S, Schneiderman JE, Tein I. L-Arginine Affects Aerobic Capacity and Muscle Metabolism in MELAS (Mitochondrial Encephalomyopathy, Lactic Acidosis and Stroke-Like Episodes) Syndrome. *PloS one* 2015;10:e0127066.
60. Koene S, Smeitink J. Mitochondrial medicine: entering the era of treatment. *Journal of internal medicine* 2009;265:193-209.
61. Rodenburg RJ. Biochemical diagnosis of mitochondrial disorders. *Journal of inherited metabolic disease*.
62. Tachibana M, Sparman M, Sritanandomchai H, *et al*. Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature* 2009;461:367-72.
63. Viscomi C, Bottani E, Zeviani M. Emerging concepts in the therapy of mitochondrial disease. *Biochimica et biophysica acta* 2015;1847:544-57.
64. Viscomi C, Bottani E, Civiletto G, *et al*. In vivo correction of COX deficiency by activation of the AMPK/PGC-1alpha axis. *Cell metabolism* 2011;14:80-90.
65. Khan NA, Auranen M, Paetau I, *et al*. Effective treatment of mitochondrial myopathy by nicotinamide riboside, a vitamin B3. *EMBO molecular medicine* 2014;6:721-31.
66. Liu S, Jing F, Yu C, Gao L, Qin Y, Zhao J. AICAR-Induced Activation of AMPK Inhibits TSH/SREBP-2/HMGCR Pathway in Liver. *PloS one* 2015;10:e0124951.
67. Brass EP, Hoppel CL. Mitochondria as targets of drug toxicity: Lessons from the R118 Phase 1 experience. *Clinical Pharmacology & Therapeutics* 2015.
68. Johnson SC, Yanos ME, Kayser EB, *et al*. mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. *Science* 2013;342:1524-8.
69. Wenz T, Luca C, Torraco A, Moraes CT. mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription. *Cell metabolism* 2009;9:499-511.
70. Ahola-Erkila S, Carroll CJ, Peltola-Mjosund K, *et al*. Ketogenic diet slows down mitochondrial myopathy progression in mice. *Human molecular genetics* 2010;19:1974-84.
71. Ahola S, Auranen M, Isohanni P, *et al*. High-fat diet results in targeted death of ragged-red fibers in mitochondrial myopathy. *Euromit*; 2014; Tampere.
72. Koene S, Smeitink J. Metabolic manipulators: a well founded strategy to combat mitochondrial dysfunction. *Journal of inherited metabolic disease*.
73. Koopman WJ, Nijtmans LG, Dieteren CE, *et al*. Mammalian mitochondrial complex I: Biogenesis, Regulation and Reactive Oxygen Species generation. *Antioxid Redox Signal* 2009.
74. Balaban RS, Nemoto S, Finkel T. Mitochondria, oxidants, and aging. *Cell* 2005;120:483-95.
75. Murphy MP. How mitochondria produce reactive oxygen species. *The Biochemical journal* 2009;417:1-13.
76. Koopman WJ, Verkaart S, Visch HJ, *et al*. Human NADH:ubiquinone oxidoreductase deficiency: radical changes in mitochondrial morphology? *American journal of physiology* 2007;293:C22-9.
77. Koopman WJ, Distelmaier F, Hink MA, *et al*. Inherited complex I deficiency is associated with faster protein diffusion in the matrix of moving mitochondria. *American journal of physiology* 2008;294:C1124-32.
78. Lenaz G, Genova ML. Mobility and function of coenzyme Q (ubiquinone) in the mitochondrial respiratory chain. *Biochimica et biophysica acta* 2009;1787:563-73.
79. Glover EI, Martin J, Maher A, Thornhill RE, Moran GR, Tarnopolsky MA. A randomised trial of coenzyme Q10 in mitochondrial disorders. *Muscle & nerve* 2010;42:739-48.
80. Mashima Y, Hiida Y, Oguchi Y. Remission of Leber's hereditary optic neuropathy with idebenone. *Lancet* 1992;340:368-9.

81. Mashima Y, Kigasawa K, Wakakura M, Oguchi Y. Do idebenone and vitamin therapy shorten the time to achieve visual recovery in Leber hereditary optic neuropathy? *J Neuroophthalmol* 2000;20:166-70.
82. Di Prospero NA, Baker A, Jeffries N, Fischbeck KH. Neurological effects of high-dose idebenone in patients with Friedreich's ataxia: a randomised, placebo-controlled trial. *Lancet Neurol* 2007;6:878-86.
83. Meier T, Buyse G. Idebenone: an emerging therapy for Friedreich ataxia. *J Neurol* 2009;256 Suppl 1:25-30.
84. Martinelli D, Catteruccia M, Piemonte F, *et al.* EPI-743 reverses the progression of the pediatric mitochondrial disease-Genetically defined Leigh Syndrome. *Molecular genetics and metabolism* 2012.
85. Sadun AA, Chicani CF, Ross-Cisneros FN, *et al.* Effect of EPI-743 on the clinical course of the mitochondrial disease Leber hereditary optic neuropathy. *Arch Neurol* 2012;69:331-8.
86. Chicani CF, Chu ER, Miller G, Kelman SE, Sadun AA. Comparing EPI-743 treatment in siblings with Leber's hereditary optic neuropathy mt14484 mutation. *Can J Ophthalmol* 2013;48:e130-3.
87. [www.clinicaltrials.gov](http://www.clinicaltrials.gov). (Accessed 22-06, 2015,
88. Bresolin N, Doriguzzi C, Ponzetto C, *et al.* Ubidecarenone in the treatment of mitochondrial myopathies: a multi-centre double-blind trial. *J Neurol Sci* 1990;100:70-8.
89. Chinnery P, Majamaa K, Turnbull D, Thorburn D. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2006:CD004426.
90. Brunner HI, Ravelli A. Developing outcome measures for paediatric rheumatic diseases. *Best practice & research Clinical rheumatology* 2009;23:609-24.
91. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction. *J Rheumatol* 1993;20:528-30.
92. Organisation WH. *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization Press; 2001.
93. Ronen GM, Fayed N, Rosenbaum PL. Outcomes in pediatric neurology: a review of conceptual issues and recommendationsThe 2010 Ronnie Mac Keith Lecture. *Developmental medicine and child neurology* 2011;53:305-12.
94. Merlini L, Bertini E, Minetti C, *et al.* Motor function-muscle strength relationship in spinal muscular atrophy. *Muscle & nerve* 2004;29:548-52.
95. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disabil Rehabil* 2005;27:315-40.
96. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J. Issues for selection of outcome measures in stroke rehabilitation: ICF Body Functions. *Disabil Rehabil* 2005;27:191-207.
97. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF Participation. *Disabil Rehabil* 2005;27:507-28.
98. Higgins PA, Straub AJ. Understanding the error of our ways: mapping the concepts of validity and reliability. *Nursing outlook* 2006;54:23-9.
99. Vet HTCM, LB; Knol, DL. *Measurement in Medicine*. Cambridge: Cambridge University Press; 2011.





# Scope, aims and outline of this thesis





## Scope of this thesis

At the time this project was started, little attention had been paid to the methodological quality of clinical studies in mitochondrial disorders. Clinical management was mainly directed by non-blind studies in small, heterogeneous cohorts, using poorly examined and/or clinically insignificant outcome measures. In an international collaboration, we united in our combat for high-quality research in this challenging field. One of the recommendations to improve the design of future clinical trials was to use validated and clinically meaningful end points. Since experience with the validation of end points for mitochondrial disease field was lacking, this project had been started.

## Aims of this thesis

The main aim of this thesis was to contribute to the knowledge on which outcome measures to use in future trials in children with mitochondrial disorders.

This aim led to the following research goals:

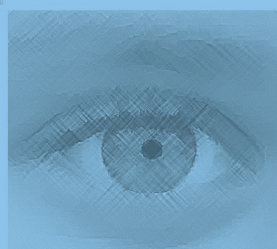
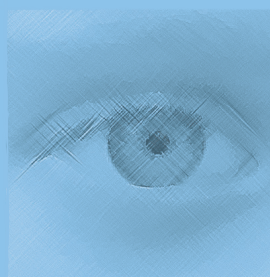
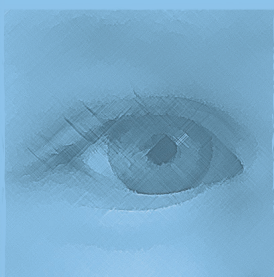
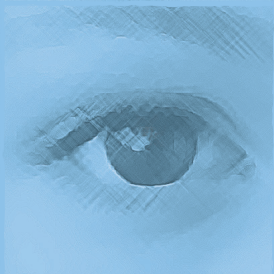
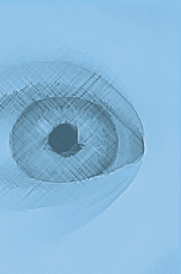
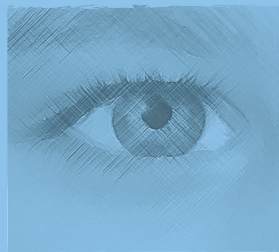
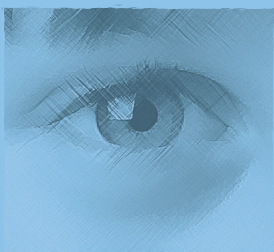
- I. Explore which the symptoms and disabilities associated with mitochondrial disease should be measured in clinical trials from a patients' and doctors' perspective.
- II. Systematically select instruments from literature that could be used to cover these symptoms and disabilities.
- III. Validate these instruments in patients with mitochondrial disorders.
- IV. Develop new instruments for domains that seem important but are not covered.

## Outline of this thesis

The first part of this thesis aims to select clinically relevant and robust outcome measures for children with mitochondrial disorders for the validation studies in part three. The first chapter evaluates the methodological quality of the treatment studies in mitochondrial disease so far. It concludes with recommendations to improve the quality of future clinical trials. In the second chapter, we describe the complaints and disabilities in mitochondrial disorders from a patient and parent perspective. The most frequent and most burdensome symptoms described in this chapter are used in the third chapter to select end points measuring these symptoms and limitations from literature. Using a systematic selection procedure to review the available information obtained in more or less similar disorders, we generate a toolbox of clinically relevant and robust outcome measures.

The second part aims at determining which selection (genetic or biochemical entity or clinical syndrome) should be used to define inclusion criteria for the validation studies. Part two assesses natural history in two patient cohorts: patients with a mutation in one of the nuclear genes encoding subunits or assembly proteins for Complex I (Chapter 4) and in carriers of the m.3243A>G mutation (Chapter 5).

The third part of this thesis describes the results of several validation studies using the instruments from the toolbox composed in chapter three. The first three studies are exploratory studies in the paediatric population. In the sixth chapter, we describe the feasibility of 3D accelerometers in measuring daily physical activity in children with mitochondrial disease. In chapter 7, we describe the feasibility, reliability and validity of the (motor-)assisted 6-minute cycling test in children with mitochondrial and neuromuscular disorders. Since the available instrument to determine general mitochondrial disease severity in children was in our opinion not suitable for clinical trials, we developed the International Paediatric Mitochondrial Disease Scale, which assesses the severity of mitochondrial disease related complaints and disabilities from a subjective, objective and functional perspective. Chapter 8 describes the process of the development, the initial reliability and validity studies and the results of the international validation study. The last three chapters of the third part describe studies in a large cohort of m.3243A>G carriers, who have been included in a national follow-up in our centre since 2010 (see Chapter 5). In the ninth chapter of this thesis, we describe the feasibility of 2-dimensional strain echocardiography, a technique to quantify the deformation of the myocardium non-invasively. The value of two potential biomarkers (fibroblast growth factor 21 (FGF21) and growth and differentiation factor 15 (GDF15)) in determining and monitoring disease progression in m.3243A>G carriers is studied in Chapter 10 and 11.



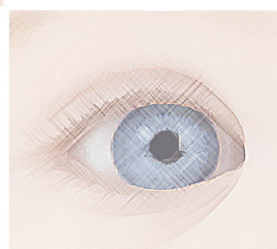
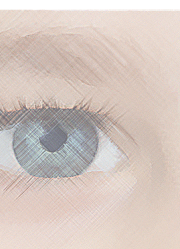
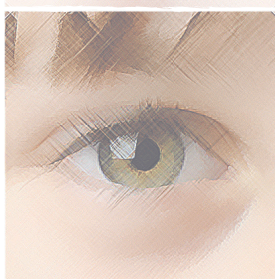
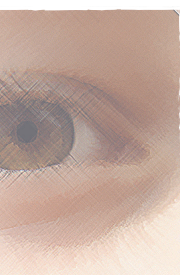
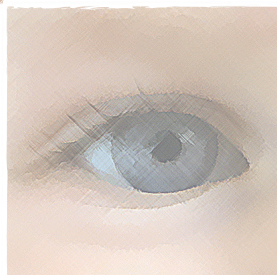
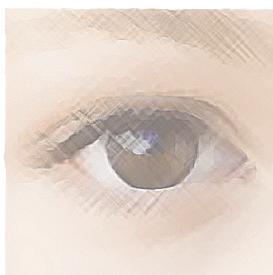
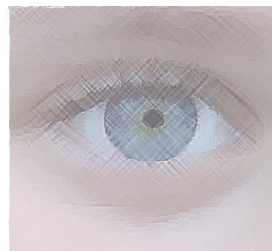
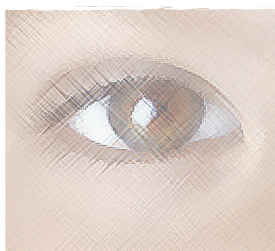
# Part I

Studying interventions in  
mitochondrial disorders:  
past, present and future



## Goal

To select clinically relevant and robust outcome measures for in children with mitochondrial disease from literature.





# Chapter 1

New treatments for  
mitochondrial disease - no  
time to drop our standards



## New treatments for mitochondrial disease - no time to drop our standards

Gerald Pfeffer<sup>1,2,3,4</sup>, Rita Horvath<sup>1,2,3,4</sup>, Thomas Klopstock<sup>5,6,7,8</sup>, Vamsi K. Mootha<sup>9</sup>, Anu Suomalainen<sup>10</sup>, **Saskia Koene**<sup>11</sup>, Michio Hirano<sup>12</sup>, Massimo Zeviani<sup>13</sup>, Laurence A. Bindoff<sup>14</sup>, Patrick Yu Wai Man<sup>1,2,3,4</sup>, Michael Hanna<sup>3,15</sup>, Valerio Carelli<sup>16,17</sup>, Robert McFarland<sup>1,3,4,18</sup>, Kari Majamaa<sup>19,20</sup>, Douglass M. Turnbull<sup>1,3,4,17</sup>, Jan Smeitink<sup>11</sup>, Patrick F. Chinnery<sup>1,2,3,4</sup>

<sup>1</sup> Wellcome Trust Centre for Mitochondrial Research, Newcastle University, Newcastle upon Tyne, NE1 3BZ, United Kingdom (UK)

<sup>2</sup> Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, NE1 3BZ, UK

<sup>3</sup> Medical Research Council Centre for Translational Research in Neuromuscular Diseases, UK

<sup>4</sup> Royal Victoria Infirmary, Newcastle upon Tyne Hospitals NHS Foundation Trust, NE1 4LP, UK

<sup>5</sup> Dept. of Neurology, Friedrich-Baur-Institute, Ludwig-Maximilians-University, Munich, Germany

<sup>6</sup> German Network for mitochondrial disorders (mitoNET)

<sup>7</sup> German Centre for Vertigo and Balance Disorders, Munich, Germany

<sup>8</sup> DZNE - German Centre for Neurodegenerative Diseases, Munich, Germany

<sup>9</sup> Department of Molecular Biology, Centre for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

<sup>10</sup> Biomedicum-Helsinki, r.C512A, Haartmaninkatu 8, 00290 Helsinki, Finland

<sup>11</sup> Nijmegen Centre for Mitochondrial Disorders, Radboudumc, Netherlands

<sup>12</sup> Department of Neurology, Columbia University Medical Centre, 630 West 168th Street, New York, New York 10032, USA

<sup>13</sup> Unit of Molecular Neurogenetics, The Foundation 'Carlo Besta' Institute of Neurology IRCCS, Milan, Italy

<sup>14</sup> University of Bergen, IKM, & Department of Neurology, Haukeland University Hospital, Bergen, 5021, Norway

<sup>15</sup> UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK

<sup>16</sup> IRCCS Istituto delle Scienze Neurologiche, Bologna, Italy

<sup>17</sup> Dipartimento di Scienze Biomediche e NeuroMotorie (DIBINEM), University of Bologna, Bologna, Italy

<sup>18</sup> Institute of Ageing and Health, Newcastle University, UK

<sup>19</sup> Institute of Clinical Medicine, Department of Neurology, University of Oulu, P.O. Box 5000, FI-90014, Oulu, Finland

<sup>20</sup> Clinical Research Centre, Oulu University Hospital, University of Oulu, P.O. Box 5000, FI-90014, Oulu, Finland

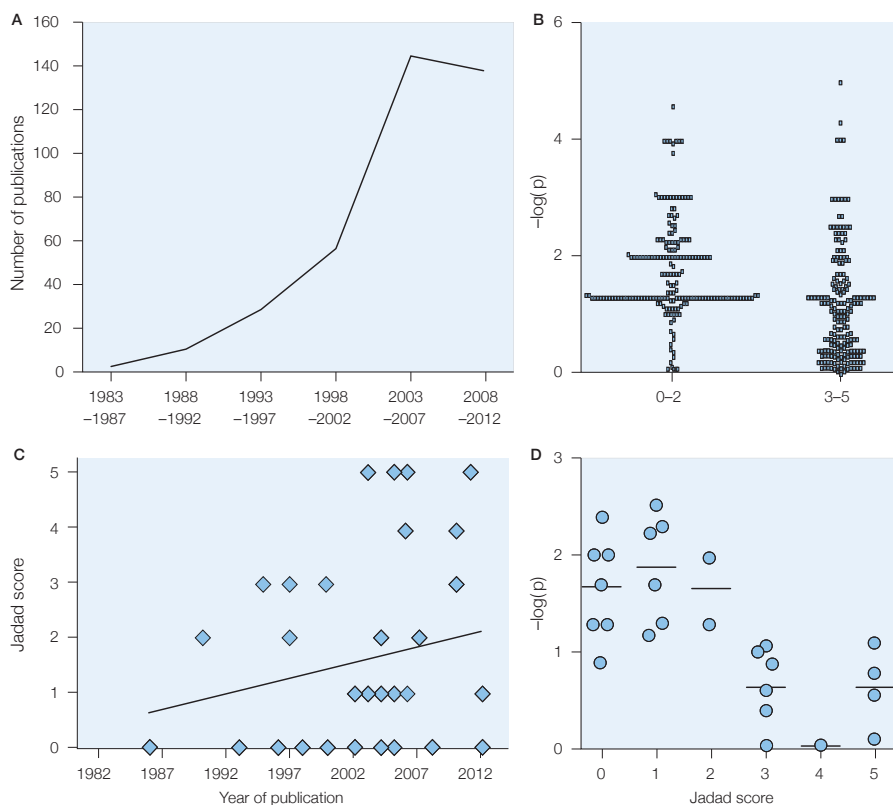
Nature Reviews Neurology 2013 Aug;9(8):474-81

Mitochondrial dysfunction is a common cause of inherited multi-system disease that often involves the nervous system. Despite major advances in our understanding of the pathophysiology of mitochondrial diseases, clinical management of these conditions remains largely supportive. Using a systematic approach, we identified 1,039 publications on treatments for mitochondrial diseases, only 35 of which included observations on more than five patients. Reports of a positive outcome on the basis of a biomarker of unproven clinical significance were more common in non-randomised and non-blinded studies, suggesting a publication bias toward positive but poorly executed studies. Although trial design is improving, there is a critical need to develop new biomarkers of mitochondrial disease. In this article, we make recommendations for the design of future treatment trials in mitochondrial diseases. Patients and physicians should no longer rely on potentially biased data, with the associated costs and risks.

## Introduction

Rare diseases affect 7% of the population and, owing to neurological involvement, many individuals with such disorders will present to a neurologist.<sup>1</sup> The vast majority of rare diseases are genetically inherited, and most have no available treatment, as identification of effective therapeutic agents for rare diseases is a difficult process. The patient population for rare disorders is usually small and distributed over a wide geographical area, often crossing administrative boundaries. Such factors limit studies of natural history, and can hinder the identification of appropriate, clinically relevant and validated disease end points. Given that the target population in rare disease is small, financial incentives for pharmaceutical companies to develop and test novel treatments are lacking. Thankfully, this lack of incentive is mitigated by legislation and national plans for such diseases,<sup>2</sup> which have kindled increasing interest from pharmaceutical corporations in these niche areas. Interest in therapeutic research in rare diseases is also driven by the hope that medicines for 'orphan diseases' might be useful for more common ailments.

Mitochondrial disorders - a group of rare inherited diseases of energy metabolism - often present with neurological features, and provide an excellent illustration of the problems associated with treatment development for rare diseases. Despite an increase over the past two decades in the number of published studies reporting treatment effects in mitochondrial disease (Figure 1A), a recent systematic review of the literature found no evidence of an effective intervention for any mitochondrial disorder.<sup>3</sup> Thanks to advances in molecular diagnostics, however, a growing number of patients with mitochondrial disorders are being identified, and the pressure to find a cure has consequently continued to mount. Mitochondrial disorders are now considered among the most common inherited diseases and, given their relentlessly progressive nature, often worsening over many decades, these disorders cause substantial morbidity. Although mitochondrial disorders can be caused by many different genetic defects of both nuclear and mitochondrial DNA (mtDNA), they share common pathogenic pathways that are potentially amenable to intervention. Here, we critically evaluate proposed treatments for mitochondrial diseases, highlighting the danger of relying on openlabel studies, and making recommendations for future trials aimed at developing new therapies for these devastating diseases.



**Figure 1.**

**Trials of treatments for mitochondrial disease.** A) Publications listed on MEDLINE in 5 year intervals show that the number of trials has increased over time (the dip at 2008–2012 is probably attributable to ascertainment before the end of 2012 on an exponential curve; see Box 1 for search and methodological details); B) Scatter plot of the negative log of all  $p$  values listed in included studies (higher numbers indicate a more statistically significant result,  $-\log(p) > 1.3 = p < 0.05$ ). Lower-quality studies had higher reported statistical significances; C) Scatter plot and trend line show improvement of study quality over time; D) Scatter plot of the negative log of all clinically relevant  $p$  values in the included studies. Lower-quality studies report greater statistical significance for these end points, which were all non-significant in high-quality studies.

## Proposed treatments

Mitochondrial disorders are primarily due to a biochemical defect of ATP synthesis. ATP is required for all active cellular processes, and the majority is generated by mitochondrial oxidative phosphorylation (OXPHOS), which facilitates the transfer of electrons between the respiratory chain enzyme complexes. For the most part, early attempts to develop treatments for mitochondrial disorders have focused on enhancing respiratory chain function (Table 1).

**Table 1.**  
**Treatments evaluated in patients with mitochondrial diseases.**

Agent	Specific mechanism(s) of action	Highest level of clinical study in humans	Reference
<i>Increase of substrate supply to respiratory chain</i>			
Carnitine	Fatty acid transfer for citric acid cycle intermediates	Case report	71
Niacin	Precursor for NADH, which transfers electrons from intermediates to the respiratory chain	Case report	72
Thiamine	Enhancement of pyruvate dehydrogenase to decarboxylate pyruvate for oxidation	Case report	73
Dichloroacetate	Inhibition of pyruvate dehydrogenase kinase to increase availability of pyruvate for oxidation	Randomised, placebo-controlled cross-over trial in MELAS due to m.3243A>G mutation (negative outcome)	34
<i>Augmentation of respiratory chain components</i>			
Riboflavin	Precursor for flavin adenine dinucleotide, an electron carrier bound to complexes I and II	Open-label study in complex I deficiency (positive outcome)	8
Coenzyme Q10	Electron carrier from complexes I and II to complex III	Randomised, placebo-controlled cross-over trial (negative outcome)	32
Idebenone	Analogue of coenzyme Q10	Randomised, placebo-controlled trial in Leber hereditary optic neuropathy (negative outcome)	36
EPI-743	Analogue of vitamin E	Open-label study in Leigh syndrome and Leber hereditary optic neuropathy (positive outcome)	39,47
<i>Bypass of respiratory chain components</i>			
Succinate	Citric acid cycle intermediate which donates electrons directly to complexes I and II, thus partially bypassing complex I	Case report	12
Vitamins C and K	Bypass of complex III	Case report	13
<i>Energy buffering</i>			
Creatine	ATP storage in muscles via the creatine phosphokinase system	Randomised, placebo-controlled cross-over trials in mitochondrial myopathies (negative outcomes in two trials, positive surrogate end points in one trial)	35,52,73

<i>Antioxidant activity</i>			
Cysteine	Increases muscle availability of glutathione peroxidase	Randomised, placebo-controlled cross-over trial in progressive external ophthalmoplegia (negative outcome)	14
Lipoic acid	$\beta$ -ketoacid dehydrogenase cofactor with antioxidant properties	Case report; randomised, placebo-controlled cross-over trial (with creatine and coenzyme Q10; negative outcomes in various mitochondrial myopathies)	66,46
Dimethylglycine	Antioxidant activity	Randomised, placebo-controlled cross-over trial in Saguenay Lac-St-Jean cytochrome c oxidase deficiency (negative outcome)	15
<i>Oxidative capacity adaptations</i>			
Aerobic exercise training	Reversal of deconditioning and/or mitochondrial adaptation to improve oxidative capacity	Randomised, non-blinded controlled trial in mitochondrial myopathies (positive outcome)	17,51,65,75
Resistance exercise training	Myofibre regeneration and presumed gene shifting	Open-label study (positive outcome)	76,77
Nitric oxide metabolism			
L-arginine	Substrate for nitric oxide synthase	Open-label placebo-controlled trial in MELAS due to m.3243A>G mutation (positive outcome)	18

Supplements aimed at increasing respiratory chain substrate availability include carnitine (which facilitates the transfer of fatty acids, thereby increasing the availability of metabolites from the citric acid cycle),<sup>4</sup> niacin (the precursor to NADH, which transfers electrons from intermediate metabolites to the respiratory chain),<sup>5</sup> and thiamine (which enhances pyruvate dehydrogenase activity and, therefore, the availability of decarboxylase pyruvate for oxidation).<sup>6</sup> A synthetic agent, dichloroacetate - an inhibitor of pyruvate dehydrogenase kinase - has also been used for treatment of mitochondrial disorders on the rationale that this compound increases the availability of pyruvate for oxidation.<sup>7</sup>

Attempts to enhance electron transfer within the respiratory chain have included supplementation with riboflavin (the precursor for flavin adenine dinucleotide (FAD), an electron carrier bound to complexes I and II),<sup>8</sup> and coenzyme Q10 (CoQ<sub>10</sub>, also known as ubiquinone, which is an electron carrier from Complexes I and II to Complex III).<sup>9</sup> Synthetic agents based on CoQ<sub>10</sub> and vitamin E - such as the drugs idebenone and EPI 743 - have also been designed to increase the penetration of an electron carrier into mitochondria and/or central nervous system tissue.<sup>10,11</sup>

Alternative strategies to treat mitochondrial diseases include biochemical 'bypass' of specific respiratory chain complexes, such as with the use of succinate (a citric



acid cycle intermediate that donates electrons directly to FAD, thus partially bypassing Complex I)<sup>12</sup> and a combination of vitamins C and K (in order to bypass Complex III).<sup>13</sup> Other treatments have focused on the reduction of toxic metabolites through antioxidant activity, and specific agents with this effect include cysteine, vitamins C and E, lipoic acid, and dimethylglycine.<sup>14,15</sup> Another approach is ‘energy buffering’; that is, the use of creatine to increase ATP storage through the creatine phosphokinase system.<sup>16</sup> Finally, exercise therapy is thought to produce adaptations in mitochondria that improve oxidative capacity and/or reduce muscle deconditioning.<sup>17</sup> Exceptions to the above categories include the use of L-arginine in patients with stroke-like episodes (in light of the vasoactive effects of this compound that are mediated through the nitric oxide pathway),<sup>18</sup> and corticosteroids.<sup>19</sup> Several other experimental treatments are in the preclinical phase of development, and have not been tried in patients to date.<sup>20</sup>

Although the first case report of a treatment benefit in mitochondrial disease was published in 1981, the first trial was not published until 1990,<sup>21</sup> and the vast majority of proposed therapies have not been tested in controlled trials. Not surprisingly, both patients and physicians are desperate to find any treatment that helps and, in the absence of hardcore evidence, clinical practice continues to be shaped by studies that involve fewer than five patients - often anecdotal evidence and case reports. Despite lack of proven efficacy, many ‘traditional’ treatments (such as CoQ<sub>10</sub>, thiamine and carnitine) are used widely,<sup>22</sup> in part owing to the low incidence of adverse effects with these therapies. After a prolonged period with no new therapies, however, recent results from openlabel studies of new agents have generated interest from patients and patient support groups.<sup>23,24</sup> Given the inherent difficulties of conducting randomised clinical trials for rare diseases, should we settle for these openlabel data?

## Reliability of evidence

To address the reliability of current evidence of efficacy for mitochondrial therapies, we objectively evaluated all of the published data on treatments for mitochondrial disease. Our aim was to determine whether less rigorous studies (that is, non-randomised, non-blinded studies) can reliably inform clinical decision-making in mitochondrial medicine. A systematic review, performed on 23rd October 2012, yielded 1,039 publications spanning a 47-year period (Box 1). Titles and abstracts were reviewed to include only studies describing treatment effects in mitochondrial diseases in five or more patients, which led to identification of a total of 35 studies.<sup>8,14,15,17,18,21,25–52</sup> The methodological quality of each study was independently evaluated by three authors using the Jadad scale (Box 2, Supplementary Document 1).<sup>53</sup> Studies are awarded a score on this scale on the basis of three factors: randomisation (up to 2 points if a valid randomisation procedure was specified), blinding (up to 2 points if a valid blinding

procedure was specified), and participant-withdrawal characteristics (1 point if withdrawals were correctly documented). The final score ranges from 0-5, with high values denoting good-quality studies and lower scores indicating poor-quality studies.

### **Box 1: Systematic review methods.**

We identified studies of English-language publications on MEDLINE using OvidSP via the following searches: 'mitochondrial disease OR mitochondrial disorder OR mitochondrial myopathy OR Leber optic neuropathy OR Leber optic neuropathy OR Leigh syndrome OR Leigh syndrome OR congenital lactic acidosis OR progressive external ophthalmoplegia OR Kearns Sayre syndrome OR mitochondrial encephalomyopathy lactic acidosis and stroke-like episodes (MELAS)' and 'treatment OR therapy OR coenzyme Q<sub>10</sub> OR idebenone OR EPI 743 OR creatine OR carnitine OR vitamin OR exercise OR arginine OR dichloroacetate'. Our search, performed on 23rd October 2012, yielded 1,039 results. Titles and abstracts were reviewed, and 49 original research studies in humans that tested a therapeutic agent in patients with mitochondrial disorders were identified. After exclusion of studies involving fewer than five patients, a total of 35 studies remained; the full-length articles for these articles were reviewed in detail. For trials with a single primary end point, these end points were selected as the relevant end point: only three trials met this criterion.<sup>15,34,36</sup> For all other trials, which had multiple end points, we selected end points that involved clinically important measures: in the case of multi-system disorders, these end points included quality of life measurements, combined scores such as the GATE or Newcastle scores, or measurements that clearly indicated relevance to patient symptomatology. In the case of mitochondrial myopathies, we included global measures of muscle strength (Medical Research Council scales), functional muscle tests (walking tests) or other standardised neurological examination results as applicable. In the case of Leber hereditary optic neuropathy, we included improvement in visual acuity as the relevant end point. In the case of one study,<sup>2</sup> the *p* value was not provided in the manuscript for treatment effect, and this was calculated ourselves using a grouped two-tailed Student *t*-test.

**Box 2: Jadad scoring.** The Jadad scores (minimum score 0, maximum score 5) for identified studies were determined by three independent reviewers. When authors disagreed the studies were re-reviewed and discussed until the most appropriate scoring was agreed upon. Points were allocated as follows:

#### **Randomisation**

- +1 if study described as randomised
- 1 if an inappropriate method of randomisation was described
- +1 if an appropriate method of randomisation was described

#### **Blinding**

- +1 if the study was described as double-blind
- 1 if an inappropriate method of blinding was described
- +1 if an appropriate method of double-blinding was described

#### **Withdrawals**

- +1 if participant withdrawals were accounted for

## Study trends

On the basis of our analysis, several trends with regard to the studies on mitochondrial treatment were observed. First, non-randomised and non-blinded studies were substantially more likely to report statistically significant results with lower  $p$  values (that is, a higher level of significance) than were randomised and blinded studies (Figure 1B). Notably, a trend towards improved study design has been observed over the past decade (Figure 1C). Second, studies with a low Jadad score were more likely to report a clinically relevant, statistically significant outcome, whereas none of the clinically-relevant primary end points (Supplementary Document 2) were statistically significant in high-quality studies (Figure 1D). The inevitable subjectivity of many direct clinical measures (such as muscle strength), together with the well-recognized placebo effect, can often account for positive results in clinical trials. These two factors are particularly problematic in open-label (non-blinded) studies, making them more vulnerable to bias. Furthermore, open-label trials involving young children can reveal 'improvements' in outcome that are due to normal growth and development, as has been demonstrated in studies of other neuromuscular disorders.<sup>54</sup>

## Publication bias

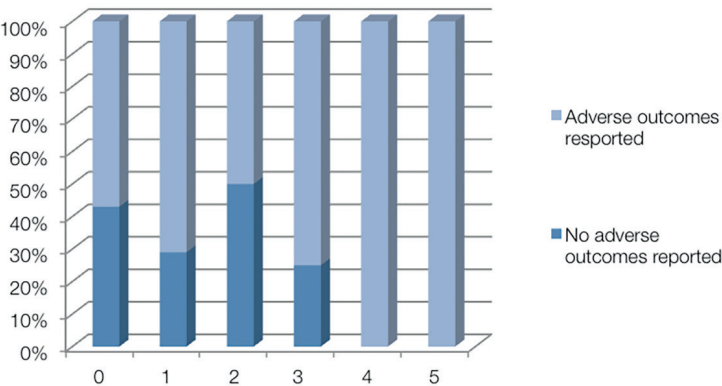
Although some treatment effects seen in open-label studies could be important, overall our findings strongly suggest a publication bias towards small, non-blinded studies that report positive effects of treatments for mitochondrial disease, despite the fact that the findings are not supported by larger randomised studies. This issue of lack of reproducibility is likely to reflect the 'winner's curse', whereby small studies that are carried out without a clearly defined end point, and without a predefined statistical analysis plan, are likely to yield a positive result, particularly if non-blinded and non-randomised. For example, CoQ<sub>10</sub> and carnitine were studied on several occasions and, in both cases, positive open-labelled studies preceded negative randomised controlled trials (most pertinently observed in a single study with both open-label and blinded phases). The positive outcome in the open-label studies was partly due to the use of non-validated surrogate disease markers in the early studies, and was compounded by a lack of blinding and/or randomisation. The higher-quality trials showed no treatment effect, despite using much higher doses of the drugs.<sup>32</sup>

In spite of these negative data, both CoQ<sub>10</sub> and carnitine continue to be prescribed in major treatment centres, and these prescriptions are then renewed indefinitely by other practitioners. Consequently, some drugs have been 'grandfathered' into prescribing practice on the basis of pre-existing positive open-label data - these drugs seem to be exempt from refutation with high-quality evidence. Paradoxically, the apparent safety of most of these agents has contributed to the problem: the balance of low risk of adverse events with a possible benefit of treatment has motivated continued

prescriptions. Hopefully, the CoQ<sub>10</sub> issue will be finally resolved, one way or another, following the publication of results from an ongoing randomised double-blind multi-centre trial of this treatment.<sup>55</sup>

### Recommendations for future trials

As well as their role in proving efficacy, large randomised trials are of critical importance with regard to patient safety. Dichloroacetate is no longer prescribed in adults with MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) because the drug caused onset or worsening of peripheral neuropathy in 17 of 19 patients, resulting in premature termination of the randomised controlled trial.<sup>34</sup> Acute neuropathy had been previously documented following dichloroacetate treatment in individual MELAS cases, but in some patients this effect was attributed to the mitochondrial disorder rather than the drug. Direct linkage of the neuropathic effect to the treatment became possible only when a large number of patients ( $n = 30$ ) were studied in a trial of high-quality design. Notably, nine of 10 trials with a high-quality design (Jadad score  $\geq 3$ ) included reports on adverse events, whereas these outcomes were reported in only 15 of 25 trials scoring  $\leq 2$  on the Jadad scale (Figure 2).



**Figure 2.**  
**Adverse outcome reporting improves with study quality.** Graph showing proportion of studies in each Jadad-score group that did and did not report adverse outcomes.

Small numbers of participants do not preclude a high-quality result, provided that the randomisation is appropriate. Selection bias during recruitment can easily contribute to misleading findings, particularly if the natural history of the disease is poorly understood. A mitochondrial disorder such as Leigh syndrome is particularly vulnerable to such bias as the severity of this disease fluctuates markedly over time in an unpredictable manner.

In studies of such cases, regression to the mean during spontaneous recovery of a patient could be misinterpreted as a positive therapeutic response, particularly when numerous end points are used, thereby increasing the chance of a false-positive result. Even for complex multi-system diseases, therefore, a trial should have a simple design, and aim to evaluate a predefined primary end point that is clinically relevant (Box 3). Other end points, such as biomarkers or physiological measurements, remain critically important to confirm that the agent is an effector of its target mechanism, but should not be used to prove clinical efficacy unless the end point is clinically relevant.

**Box 3: Recommendations for treatment trials in mitochondrial diseases.**

- Mitochondrial disorders are heterogeneous, and often have a complex multi-system phenotype that fluctuates over time in an unpredictable manner, which presents a major challenge for the design and interpretation of clinical trial data.
- Small studies should focus on patients with a similar genotype and phenotype, and ideally those at a similar stage of the disease; for rare mitochondrial disease this usually requires international collaboration.
- Simple trial designs are often the best, using validated, clinically meaningful and pre-specified primary end points. End points should be chosen that are most relevant to the genotype or phenotype in question.
- A key issue is the identification of biomarkers that are indicative of clinically relevant outcomes, which will require multi-disciplinary collaboration and patient involvement.
- Open-label trials are prone to bias through unanticipated placebo effects and subjective clinical measurements. These studies are important as a first step in evaluating treatments, but they must be considered preliminary, and should not shape routine clinical practice.
- Open-label studies should be published only if they have a small number of defined pre-specified end points and a clear predefined statistical analysis plan, and are publically registered on ClinicalTrials.gov before recruitment commences. The results of these studies should not be considered as preliminary evidence for the benefit and safety of an intervention, but merely serve as a signal to proceed with further evaluation in appropriately controlled trials.
- Large multi-centre randomised controlled trials have been carried out for mitochondrial disease, and several others are in progress. These trials establish proof of principle that data of the highest quality can be produced to underpin mitochondrial medicine, facilitated by international consortia.
- Off-licence prescription of medicines or food supplements could have value in a compassionate context, but the lack of objective efficacy should be made clear to patients and families, who should be advised that prescribing may stop if a high-quality negative trial is published.

Open trial results should be considered preliminary at best and, if positive, be followed up with a randomised study. Such a trial could be targeted to a specific phenotypic or genetic group on the basis of data from a pilot open-label study. On the other hand, negative results from potentially underpowered studies in patients with fluctuating phenotypes and genetically heterogeneous diseases may mask minor therapeutic benefits. Small treatment effects are important when no other treatments are available (Box 3).

## A novel approach

One approach to addressing the difficulties in obtaining high-quality evidence for rare disorders is exemplified by Health Canada's conditional approval to prescribe idebenone in Friedreich ataxia. Approval was provided under the condition that enhanced postmarketing surveillance took place under the Notice of Compliance with Conditions policy.<sup>56</sup> The original approval for this trial was based on evidence of treatment benefit from a single randomised controlled trial.<sup>57</sup> When no efficacy was demonstrated in the subsequent randomised trial for the primary or secondary end points,<sup>58</sup> Health Canada issued an open letter to physicians on 20th January 2010 to draw attention to the negative results. When further trials of idebenone in Friedreich ataxia were also negative<sup>59</sup> and systematic reviews concluded that there was no evidence of efficacy,<sup>60</sup> Santhera Pharmaceuticals announced on 27<sup>th</sup> February 2013 that the agent would be withdrawn from the market after consultation with Health Canada.<sup>61</sup> This system of drug approval enabled a balanced approach to the problem: a potentially valuable drug was made available at the earliest opportunity on the basis of high-quality evidence, with the possibility of later withdrawal if further trial data were not positive.

## Critical issues for the future

For the reasons described above, drug development in mitochondrial disorders has been highly problematic. Tightening of safety and efficacy standards is well-recognized to have led to an increase in the costs of developing novel agents for these disorders. Consequently, despite increasing investment in research and development, the number of drugs successfully brought to market each year continues to decrease.<sup>62</sup> To address this issue, the FDA introduced the Critical Path Initiative, which provides recommendations to help reconcile society's high safety expectations for novel drugs with the pharmaceutical industry's limited capacity to produce these treatments given the increasing costs.<sup>62</sup> A key component of these recommendations is multidisciplinary collaboration to identify biomarkers that correspond with clinical benefit and/or adverse reactions, in order to identify suitable or unsuitable compounds at an early phase of testing.

Mitochondrial disorders have no shortage of potential biomarkers, ranging from biochemical measurements (such as lactate, pyruvate, alanine, citrulline, creatine kinase, organic acid quantification and antioxidant levels), physiological measurements (including cardiac dimensions and/or output, visual parameters, and various measurements of aerobic or anaerobic exercise capacity or muscle power), genetic measurements (mtDNA deletion/mutation burden or copy number), and imaging (magnetic resonance spectroscopy (MRS) of brain or muscle). To date, none of the biomarkers that have been altered by treatments in high-quality studies have been shown to correlate closely with a clinical outcome (namely, lactate,<sup>23,30,32,46,49</sup> pyruvate,<sup>29</sup> alanine,<sup>29</sup> antioxidant levels,<sup>14,46</sup> and MRS findings in the brain<sup>29</sup>).

Future studies would be greatly aided by the discovery of a clinically valid biomarker or outcome measure.<sup>63,64</sup> These biomarkers may not be specific for mitochondrial disease per se; for example, a measure of cardiac, visual or auditory function could be useful in patients with an m.3243A>G mutation in the *MTTL1* gene - one of the most common mutations associated with MELAS. Agreement on an accepted biomarker would enable its use for 'screening' of new treatments in small exploratory experimental medicine studies (phase Ib or phase II), potentially revealing major adverse effects, or showing that an agent is ineffective at an early stage. Notably, a positive result from such early studies should only be used to inform planning for a randomised placebo-controlled trial (phase III), and would not provide evidence of clinical efficacy. A major pitfall of this approach is the risk that a new treatment could be rejected prematurely because it did not influence a selected biomarker (type II error, false negative). Again, this risk could be mitigated with the use of an accepted, sensitive, well-characterised biomarker that correlates with disease severity.

Finally, efficacy (phase III) studies are not without their challenges. Without detailed natural history data, it may not be possible to identify a sensitive and reliable primary trial end point that is directly related to disability. The end point may be different for each mitochondrial subphenotype, and the aim of the study will also be critical: will the treatment prevent progression (such as in Leber hereditary optic neuropathy<sup>34</sup>), reduce the frequency of relapses (for example, with L-arginine in MELAS<sup>18</sup>), or reverse a functional deficit (as in exercise studies<sup>17,51,65</sup>)? Ultimately, the aim of these studies will be to improve quality of life (QoL), but demonstrating a significant change with crude QoL questionnaires will be challenging in a study with perhaps a few hundred patients at most.

Practically speaking, recent work has shown that the most common mitochondrial syndromes are sufficiently prevalent to allow multi-centre trials to achieve adequate enrolment,<sup>66</sup> and prior trials have demonstrated effective multi-centre collaboration across multiple national jurisdictions.<sup>36,49</sup> Nevertheless, to study all treatments using this approach will not be possible. For some subgroups of mitochondrial disease, small studies will be the only way forward. We believe that such studies can be highly valuable, provided that a high-quality study design is employed (Box 3). In short, the 'big pharma' model (specifically, phase III trials) may not be possible, so other approaches should be employed if we are to make headway. The past should not be forgotten, however, and new treatments should be compared with both placebo and current best standard of care. Inevitably, this approach could involve incorporation of drugs into the trial that are already grandfathered into clinical practice, even if their adoption has a weak evidence basis.

## Conclusions

The increase in publications of trials of mitochondrial treatments over the past decade has been mirrored by a trend towards improved study design (Figure 1). These methodological advances have been underpinned by disease registries and multi-centre collaborations (such as the North American and European Mitochondrial Diseases Networks<sup>67–69</sup>), which provide proof of principle that rigorous testing of mitochondrial medicines is possible, even for this heterogeneous group of rare disorders. In general, rare disorders with incidence above five per 100,000 individuals are more likely to have orphan drugs approved for their treatment.<sup>70</sup> Cause for optimism exists, therefore, that novel treatments will continue to be trialled for mitochondrial disease.

Notably, however, premature use of a new treatment can have far-reaching consequences that are quite separate from the high cost and potential adverse effects. Ineffective medicines undermine patient confidence in both medical practitioners and the medical research community, who may be accused of exploiting patients and their families for commercial gain. Such lack of trust will blunt enthusiasm for future clinical trials. We therefore urge judicious use of off-licence medicines on a named-patient basis (also known as ‘expanded access’ or ‘special access’ programmes in North America). The hope of short-term benefit must be counterbalanced by the chance of causing longer-term damage, in part through the ‘opportunity cost’ of off-licence prescribing for a specific patient, which delays the more rigorous evaluation of newer treatments. The root cause of the problem is likely to be multi-faceted, with academics motivated by publication-linked career advancement, industry being driven by financial incentives, and patients and families driven by their immediate needs for disease improvement and health.

Resolution of these potentially conflicting issues will not be easy, but all stakeholders must work together to ensure efficient progress. Critical issues involve the identification of disease biomarkers that correspond to the clinical outcome of the patient, the use of multi-centre collaborations to include adequate patient populations for study, and multidisciplinary collaborations to identify novel agents with novel mechanisms, with innovative and accurate study design using clinically relevant primary end points. Leading mitochondrial physicians should set an example, avoiding overemphasis on the theoretical benefits of unproven treatments; patient groups can better educate their members to engage in high-quality research and controlled trials; and both should work collaboratively with industry in well-powered, multi-centre randomised controlled trials. Only by doing this will we make headway in developing treatments for these currently incurable diseases.



## References

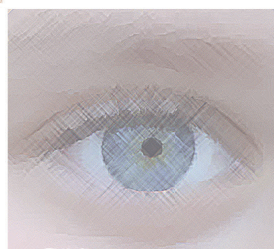
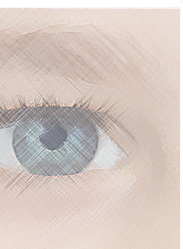
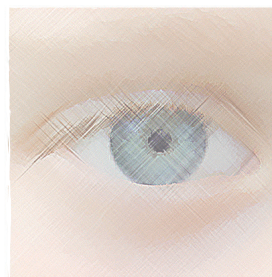
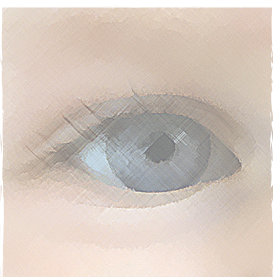
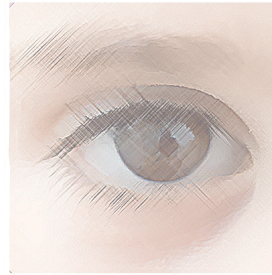
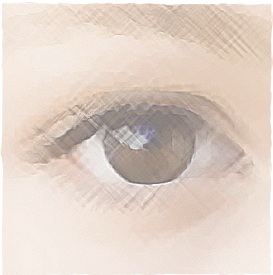
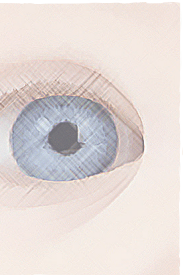
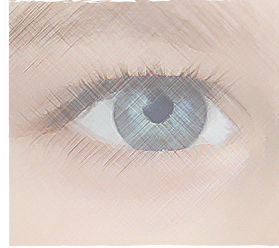
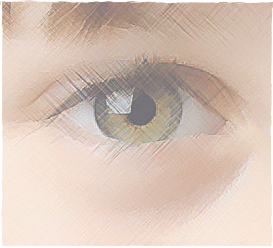
1. Murphy, S. M., Puwanant, A. & Griggs, R. C. Unintended effects of orphan product designation for rare neurological diseases. *Ann Neurol* 2012;72, 481–490.
2. Kesselheim, A. S. Ethical considerations in orphan drug approval and use. *Clin Pharmacol Ther* 2012;92, 153–155.
3. Pfeffer, G., Majamaa, K., Turnbull, D. M., Thorburn, D. & Chinnery, P. F. Treatment for mitochondrial disorders. *Cochrane Database of Systematic Reviews*, Issue 4.
4. Pons, R. & De Vivo, D. C. Primary and secondary carnitine deficiency syndromes. *J Child. Neurol* 1995;10 (Suppl. 2), S8–S24.
5. Majamaa, K., Rusanen, H., Remes, A. M., Pyhtinen, J. & Hassinen, I. E. Increase of blood NAD<sup>+</sup> and attenuation of lactic acidemia during nicotinamide treatment of a patient with the MELAS syndrome. *Life Sci* 1996;58, 691–699.
6. Lou, H. C. Correction of increased plasma pyruvate and plasma lactate levels using large doses of thiamine in patients with Kearns–Sayre syndrome. *Arch Neurol* 1981. 38, 469 .
7. Stacpoole, P. W. The pharmacology of dichloroacetate. *Metabolism* 1989;38, 1124–1144.
8. Bernsen, P. L., Gabreëls, F. J., Ruitenbeek, W. & Hamburger, H. L. Treatment of complex I deficiency with riboflavin. *J Neurol Sci* 1993; 118, 181–187.
9. Rauchová, H., Drahotá, Z. & Lenaz, G. Function of coenzyme Q in the cell: some biochemical and physiological properties. *Physiol Res* 1995; 44, 209–216.
10. Orsucci, D., Mancuso, M., Lenco, E. C., LoGerfo, A. & Siciliano, G. Targeting mitochondrial dysfunction and neurodegeneration by means of coenzyme Q10 and its analogues. *Curr Med Chem* 2011;18, 4053–4064.
11. Shrader, W. D. *et al.*  $\alpha$ -Tocotrienol quinone modulates oxidative stress response and the biochemistry of aging. *Bioorg Med Chem Lett* 2011; 21, 3693–3698.
12. Shoffner, J. M. *et al.* Spontaneous Kearns–Sayre/chronic external ophthalmoplegia plus syndromes associated with a mitochondrial DNA deletion: a slip-replication model and metabolic therapy. *Proc Natl Acad Sci USA* 1989; 86, 7952–7956.
13. Eleff, S. *et al.* <sup>31</sup>P NMR study of improvement in oxidative phosphorylation by vitamins K3 and C in a patient with a defect in electron transport at complex III in skeletal muscle. *Proc Natl Acad Sci USA* 1984 81, 3529–3533.
14. Mancuso, M. *et al.* Oxidative stress biomarkers in mitochondrial myopathies, basally and after cysteine donor supplementation. *J Neurol* 2010; 257, 774–781.
15. Liet, J. M. *et al.* The effect of short-term dimethylglycine treatment on oxygen consumption in cytochrome oxidase deficiency: a double-blind randomised cross-over clinical trial. *J Paediatr* 2003;142, 62–66.
16. Wallimann, T., Tokarska-Schlattner, M. & Schlattner, U. The creatine kinase system and pleiotropic effects of creatine. *Amino Acids* 2011;40, 1271–1296.
17. Cejudo, P. *et al.* Exercise training in mitochondrial myopathy: a randomised controlled trial. *Muscle Nerve* 2005;32, 342–350.
18. Koga, Y. *et al.* L-arginine improves the symptoms of stroke-like episodes in MELAS. *Neurology* 2005;64, 710–712.
19. Walcott, B. P. *et al.* Steroid responsive A3243G mutation MELAS: clinical and radiographic evidence for regional hyperperfusion leading to neuronal loss. *Neurologist* 2012;18, 159–170.

20. Hassani, A., Horvath, R. & Chinnery, P. F. Mitochondrial myopathies: developments in treatment. *Curr Opin Neurol* 2010; 23, 459–465.
21. Bresolin, N. *et al.* Ubidecarenone in the treatment of mitochondrial myopathies: a multi-centre double-blind trial. *J Neurol Sci* 1990;100, 70–78.
22. Kerr, D. S. Treatment of mitochondrial electron transport chain disorders: a review of clinical trials over the past decade. *Mol Genet Metab* 2010;99, 246–255.
23. CureMito [online], <http://curemito.org/> (2012).
24. United Mitochondrial Disease Foundation [online], [www.umdf.org](http://www.umdf.org) (2013).
25. Barshop, B. A. *et al.* Chronic treatment of mitochondrial disease patients with dichloroacetate. *Mol Genet. Metab* 2004;83, 138–149.
26. Blankenberg, F. G. *et al.* Brain uptake of Tc99m-HMPAO correlates with clinical response to the novel redox modulating agent EPI 743 in patients with mitochondrial disease. *Mol Genet Metab* 2012;107, 690–699.
27. Campos, Y. *et al.* Plasma carnitine insufficiency and effectiveness of L carnitine therapy in patients with mitochondrial myopathy. *Muscle Nerve* 1993;16, 150–153.
28. Chen, R. S., Huang, C. C. & Chu, N. S. Coenzyme Q10 treatment in mitochondrial encephalomyopathies. Short-term double-blind, cross-over study. *Eur Neurol* 1997;37, 212–218.
29. De Stefano, N. *et al.* Short-term dichloroacetate treatment improves indices of cerebral metabolism in patients with mitochondrial disorders. *Neurology* 1995;45, 1193–1198.
30. Duncan, G. E., Perkins, L. A., Theriaque, D. W., Neiberger, R. E. & Stacpoole, P. W. Dichloroacetate therapy attenuates the blood lactate response to submaximal exercise in patients with defects in mitochondrial energy metabolism. *J Clin Endocrinol Metab* 2004; 89, 1733–1738.
31. Enns, G. M. *et al.* Initial experience in the treatment of inherited mitochondrial disease with EPI 743. *Mol Genet Metab* 2012;105, 91–102.
32. Glover, E. I. *et al.* A randomised trial of coenzyme Q10 in mitochondrial disorders. *Muscle Nerve* 2010;42, 739–748.
33. Gold, R. *et al.* Phosphorus magnetic resonance spectroscopy in the evaluation of mitochondrial myopathies: results of a 6 month therapy study with coenzyme Q. *Eur Neurol* 1996;36, 191–196.
34. Kaufmann, P. *et al.* Dichloroacetate causes toxic neuropathy in MELAS: a randomised, controlled clinical trial. *Neurology* 2006;66, 324–330.
35. Klopstock, T. *et al.* A placebo-controlled cross-over trial of creatine in mitochondrial diseases. *Neurology* 2000;55, 1748–1751.
36. Klopstock, T. *et al.* A randomised placebo-controlled trial of idebenone in Leber's hereditary optic neuropathy. *Brain* 2011;134, 2677–2686.
37. Koga, Y. *et al.* Effects of L-arginine on the acute phase of strokes in three patients with MELAS. *Neurology* 2002;58, 827–828.
38. Komura, K., Hobbiebrunken, E., Wilichowski, E. K. & Hanefeld, F. A. Effectiveness of creatine monohydrate in mitochondrial encephalomyopathies. *Pediatr. Neurol* 2003;28, 53–58.
39. Martinelli, D. *et al.* EPI 743 reverses the progression of the paediatric mitochondrial disease-genetically defined Leigh Syndrome. *Mol Genet Metab* 2012;107, 383–388.
40. Mashima, Y., Kigasawa, K., Wakakura, M. & Oguchi, Y. Do idebenone and vitamin therapy shorten the time to achieve visual recovery in Leber hereditary optic neuropathy? *J Neuroophthalmol* 2000; 20, 166–170.

41. Mathews, P. M., Andermann, F., Silver, K., Karpati, G. & Arnold, D. L. Proton MR spectroscopic characterization of differences in regional brain metabolic abnormalities in mitochondrial encephalomyopathies. *Neurology* 1993;43, 2484–2490.
42. Mori, M., Yamagata, T., Goto, T., Saito, S. & Momoi, M. Y. Dichloroacetate treatment for mitochondrial cytopathy: long-term effects in MELAS. *Brain Dev* 2004;26, 453–458.
43. Ogawa, T. *et al.* Multi-centre trial on the early effects of silodosin on lower urinary tract symptoms associated with benign prostatic hyperplasia [Japanese]. *Hinyokika Kiyo* 2008;54, 757–764.
44. Panetta, J., Smith, L. J. & Boneh, A. Effect of high-dose vitamins, coenzyme Q and high-fat diet in paediatric patients with mitochondrial diseases. *J Inherit Metab.Dis.* 2004;27, 487–498.
45. Remes, A. M. *et al.* Ubiquinone and nicotinamide treatment of patients with the 3243A →G mtDNA mutation. *Neurology* 2002;59, 1275–1277.
46. Rodriguez, M. C. *et al.* Beneficial effects of creatine, CoQ<sub>10</sub>, and lipoic acid in mitochondrial disorders. *Muscle Nerve* 2007;35, 235–242.
47. Sadun, A. A. *et al.* Effect of EPI 743 on the clinical course of the mitochondrial disease Leber hereditary optic neuropathy. *Arch Neurol* 2012;69, 331–338.
48. Stacpoole, P. W. *et al.* Evaluation of long-term treatment of children with congenital lactic acidosis with dichloroacetate. *Paediatrics* 2008; 121, e1223–e1228.
49. Stacpoole, P. W. *et al.* Controlled clinical trial of dichloroacetate for treatment of congenital lactic acidosis in children. *Paediatrics* 2006;117, 1519–1531.
50. Suzuki, S. *et al.* The effects of coenzyme Q10 treatment on maternally inherited diabetes mellitus and deafness, and mitochondrial DNA 3243 (A to G) mutation. *Diabetologia* 1998;41, 584–588.
51. Taivassalo, T. *et al.* Endurance training and detraining in mitochondrial myopathies due to single large-scale mtDNA deletions. *Brain* 2006;29, 3391–3401.
52. Tarnopolsky, M. A., Roy, B. D. & MacDonald, J. R. Randomised control trial of creatine monohydrate in patients with mitochondrial cytopathies. *Muscle Nerve* 1997;20, 1502–1509.
53. Jadad, A. R. *et al.* Assessing the quality of reports of randomised clinical trials: is blinding necessary? *Control Clin Trials* 1996;17, 1–12.
54. Mazzone, E. *et al.* Functional changes in Duchenne muscular dystrophy: a 12-month longitudinal cohort study. *Neurology* 2011;77, 250–256.
55. US National Library of Medicine. ClinicalTrials.gov [online] 2013, <http://clinicaltrials.gov/ct2/show/NCT00432744?term=NCT00432744&rank=1>.
56. Notice of compliance with conditions—NOC/c (therapeutic products). Health Canada [online],
57. [http://www.hc-sc.gc.ca/dhp-mps/prodpharma/activit/fs-fi/nocccfs\\_accfd-eng.php](http://www.hc-sc.gc.ca/dhp-mps/prodpharma/activit/fs-fi/nocccfs_accfd-eng.php) (2013).
58. Di Prospero, N. A., Baker, A., Jeffries, N. & Fischbeck, K. H. Neurological effects of high-dose idebenone in patients with Friedreich's ataxia: a randomised, placebo-controlled trial. *Lancet Neurol* 2007;6, 878–886.
59. Lynch, D. R., Perlman, S. L. & Meier, T. A phase 3, double-blind, placebo-controlled trial of idebenone in friedreich ataxia. *Arch Neurol* 2010;67, 941–947.
60. Lagedrost, S. J. *et al.* Idebenone in Friedreich ataxia cardiomyopathy-results from a 6 month phase III study (IONIA). *Am Heart J* 2011;161, 639.e1–645.e1.
61. Kearney, M., Orrell, R. W., Fahey, M. & Pandolfo, M. Antioxidants and other pharmacological treatments for Friedreich ataxia. *Cochrane Database of Systematic Reviews*, Issue 4. Art. No.: CD007791.

62. CATENA® (idebenone)—voluntary withdrawal of CATENA® from the Canadian market. Health Canada [online],. <http://healthy Canadians.gc.ca/recall-alert-rappel-avis/hc-sc/2013/23509a-eng.php> (2013).
63. Woodcock, J. & Woosley, R. The FDA critical path initiative and its influence on new drug development. *Annu Rev Med* 2008;59, 1–12.
64. Koene, S. *et al.* Developing outcome measures for paediatric mitochondrial disorders: which complaints and limitations are most burdensome to patients and their parents? *Mitochondrion* 2013;13, 15–24.
65. Jeppesen, T. D. *et al.* Aerobic training is safe and improves exercise capacity in patients with mitochondrial myopathy. *Brain* 2006;129, 3402–3412.
66. Joppi, R., Bertele, V. & Garattini, S. Orphan drug development is not taking off. *Br J Clin Pharmacol* 2009;67, 494–502 .
67. North American Mitochondrial Disease Consortium [online], <http://rarediseasesnetwork.epi.usf.edu/NAMDC/> (2013).
68. The Children's Mitochondrial Disease Network [online], <http://www.emdn-mitonet.co.uk/> (2013).
69. mitoNET—German Network for Mitochondrial Disorders [online], <http://mitonet.org/> (2013).
70. Heemstra, H. E., van Weely, S., Büller, H. A., Leufkens, H. G. & de Vrieh, R. L. Translation of rare disease research into orphan drug development: disease matters. *Drug Discov Today* 2009;14, 1166–1173.
71. Hsu, C. C. *et al.* CPEO and carnitine deficiency overlapping in MELAS syndrome. *Acta Neurol Scand* 1995;92, 252–255.
72. Majamaa, K., Rusanen, H., Remes, A. & Hassinen, I. E. Metabolic interventions against complex I deficiency in MELAS syndrome. *Mol Cell Biochem* 1997;174, 291–296.
73. Kornblum, C. *et al.* Creatine has no beneficial effect on skeletal muscle energy metabolism in patients with single mitochondrial DNA deletions: a placebo-controlled, double-blind 31P-MRS cross-over study. *Eur J Neurol* 2005;12, 300–309.
74. Barbiroli, B. *et al.* Lipoic (thioctic) acid increases brain energy availability and skeletal muscle performance as shown by in vivo 31P-MRS in a patient with mitochondrial cytopathy. *J Neurol* 1995;242, 472–477.
75. Taivassalo, T. *et al.* Aerobic training benefits patients with mitochondrial myopathies more than other chronic myopathies. *Neurology* 1997;48, A214.
76. Fu, K. *et al.* A novel heteroplasmic tRNA<sup>Leu</sup>(UUR) mtDNA point mutation in a sporadic patient with mitochondrial encephalomyopathy segregates rapidly in muscle and suggests an approach to therapy. *Hum Mol Genet* 1996;5, 1835–1840.
77. Clark, K. *et al.* Correction of a mitochondrial DNA defect in human skeletal muscle. *Nat Genet* 1997;16, 222–224.





# Chapter 2

Developing outcome measures for paediatric mitochondrial disorders: Which complaints and limitations are most burdensome to patients and their parents?





## Developing outcome measures for paediatric mitochondrial disorders: Which complaints and limitations are most burdensome to patients and their parents?

**Saskia Koene**<sup>1</sup>, Saskia B. Wortmann<sup>1</sup>, Maaïke C. de Vries<sup>1</sup>, An I. Jonckheere<sup>1</sup>, Eva Morava<sup>1</sup>, Imelda J.M. de Groot<sup>1,2</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc

<sup>2</sup>Department of Rehabilitation, Radboudumc

Mitochondrion 2013 Jan;13(1):15-24

Since some drug intervention effects are only experienced by the patient, organisations such as the Food and Drug Administration prefer clinically meaningful outcome measures. Here, we evaluated which symptoms and limitations in daily life are most burdensome to paediatric patients with mitochondrial disorders and their parents, using two questionnaires. In a study of 78 patients, the most burdensome complaints included fatigue, behaviour and speech disturbances, epilepsy, and muscle weakness and a high degree of limitations in daily activities was found. Importantly, there was a discrepancy between what symptoms metabolic paediatricians estimated would be most burdensome compared to the patients'/caretakers' opinion. To include feasible and relevant outcome measures in intervention studies, the experience and opinions of patients and caretakers should therefore be heard.

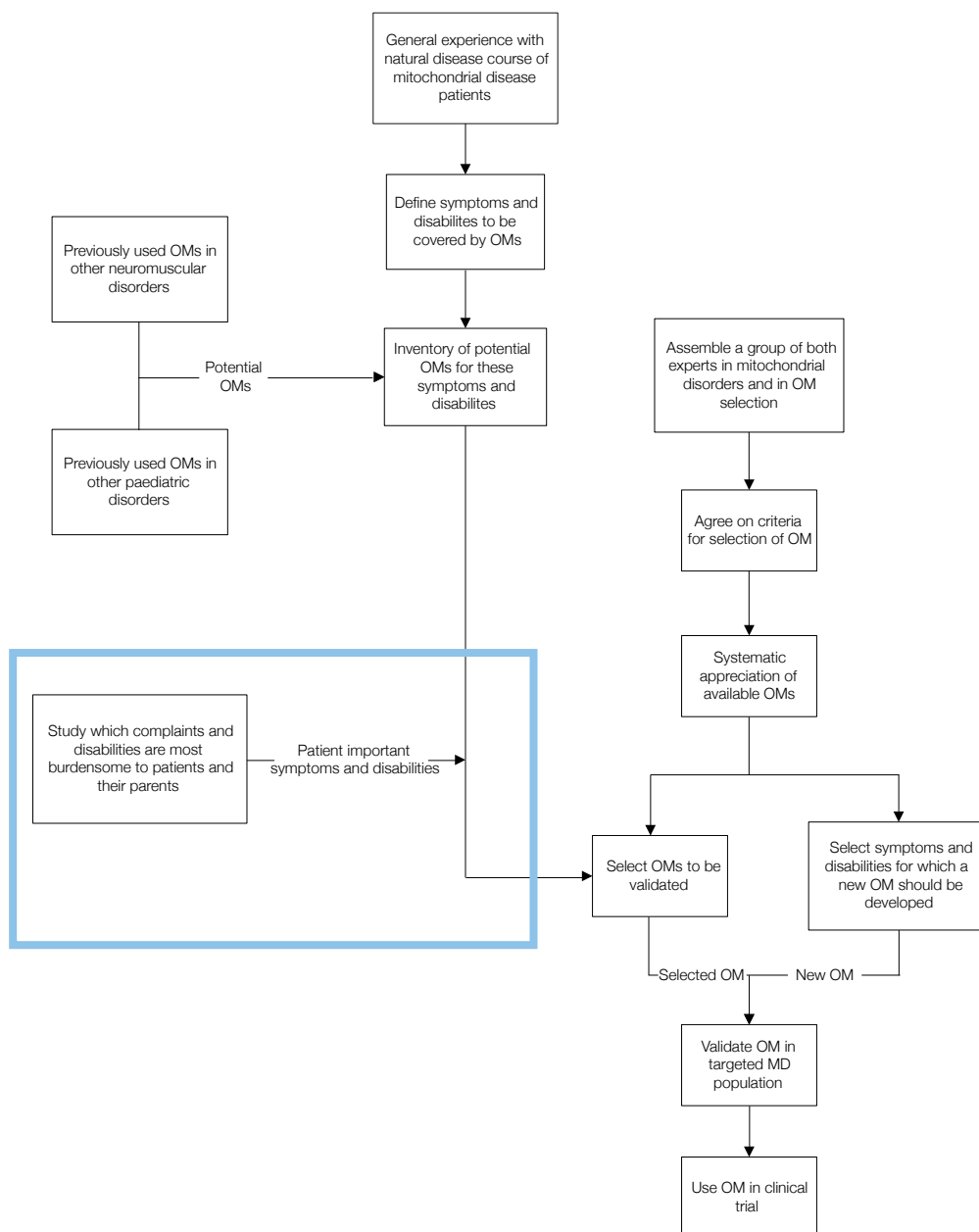
## Introduction

Mitochondrial diseases are the most prevalent inherited metabolic diseases, with an incidence of approximately 1:5,000 live births.<sup>1</sup> Currently, there is no cure for mitochondrial diseases but there are some promising results of pharmacological interventions in cells and animals<sup>2-5</sup> and in some mitochondrial disorders.<sup>6,7</sup> A recent Cochrane review states that while the number of trials in patients with mitochondrial disorders has doubled since 2006,<sup>8</sup> the percentage of trials with an adequate trial design remained stable but extremely low (12 out of 1,335 studies).<sup>9</sup> Therefore, more awareness on how to perform scientifically sound clinical trials and joint ventures to find solutions for the many difficulties associated with these heterogeneous and complex disorders is of major importance.

The careful and systematic selection of outcome measures is necessary for scientifically sound clinical trials (Figure 1). The US Food and Drug Administration (FDA) stresses that it is required to include outcome measures of importance to patients in every trial. Since the doctors' opinion of the severity of symptoms and disabilities does not necessarily agree with that of the patients,<sup>10-12</sup> we studied which symptoms and limitations are the most burdensome to paediatric patients with mitochondrial disorders and their parents and whether this corresponds to the estimation of their doctors' opinion.

Mitochondrial diseases are complex and heterogeneous multi-system disorders mostly affecting the function and sometimes the structure of the brain, muscles and heart.<sup>13</sup> Because of the complexity of these disorders, it seems difficult to determine which symptoms to measure during a clinical trial. We asked patients and their parents about the symptoms that they experience as most bothersome and most wanted to change during future treatment. To identify and classify symptoms and disabilities we used the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) framework (Figure 2),<sup>14</sup> developed by the World Health Organization (WHO). The ICF-CY checklist was used to inventory symptoms and disabilities in our patient population.

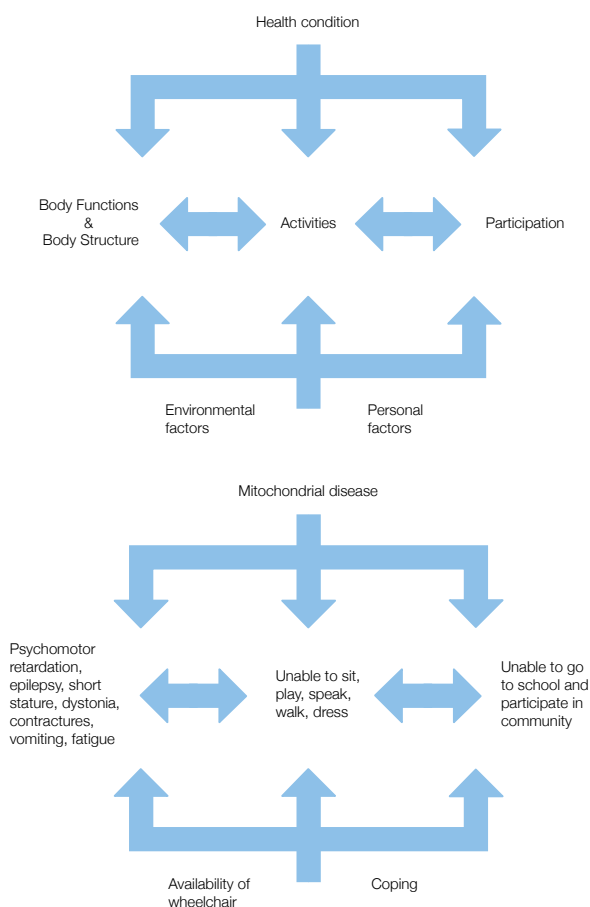
This is, to our knowledge, the first study investigating which complaints and disabilities patients with mitochondrial disorders and their parents experience in daily life.



**Figure 1.**

**Flowchart for selecting outcome measure instruments used to determine disease severity and follow-up disease progression to be used in clinical trials for children with mitochondrial disorders.**

In the red box marked the position of this study in the flowchart. MD = mitochondrial disease; OM = outcome measure.



**Figure 2.**

**The International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) framework.** The structure of the ICF-CY (left), filled out as an example for the symptoms, limitations and disabilities that can be observed in patients with mitochondrial disease (right). Note the interplay between the health condition on physical function, activity and participation on one side, and the influence of environmental and personal factors on the other.

## Methods

All paediatric (<18 years) Dutch-speaking patients with a mitochondrial disorder, followed up regularly at the Nijmegen Centre for Mitochondrial Disorders, were included in this study. Patients were defined having a mitochondrial disorder when the fresh muscle biopsy showed one or more of the following: i) ATP production 10% under the lower limit of the control range;<sup>15</sup> and/or ii) one or more enzyme complex deficiencies (under the limit of the control range); and/or iii) a confirmed pathogenic mitochon-

drial DNA (mtDNA) or nuclear mutation; and/or iv) a mitochondrial syndrome (Leigh or MELAS syndrome, MNGIE, etc.).<sup>5</sup> Patients were grouped as having myopathy, encephalomyopathy, encephalopathy or mainly gastrointestinal involvement based on their most prominent symptom(s), if they did not fit the profile of a specific mitochondrial syndrome.

We sent two questionnaires to all patients who met the inclusion criteria (Supplementary Document 3). The first questionnaire was designed to assess which symptoms were most burdensome to patients and their parents. It contained three domains: i) The presence of symptoms/complaints in the child; ii) Which three symptoms would (the parent expect) the child most like to change mostly; and iii) Which three symptoms would the parents themselves like to change? The symptoms and complaints list was composed of the most frequently observed symptoms in mitochondrial disease (literature search and subsequent agreement SK and JS). The items are listed in Figure 5. All five (trainee-) metabolic paediatricians involved in the care of these patients with mitochondrial disorders at the Nijmegen Centre for Mitochondrial Disorders were asked to fill out the first questionnaire to assess which symptoms they felt are most burdensome to patients and their parents. The questionnaire was completed by every physician (both for the parents' and for the patients' perspective). They answered, keeping in mind the whole spectrum of mitochondrial disease patients seen in our centre, the questions: i) 'Which three symptoms would (the parent expect) the child most like to change?' (patient and/or parent report); and ii) 'Which three symptoms would the parents themselves like to change (parent report)?' using the same list of symptoms patients/parents used.

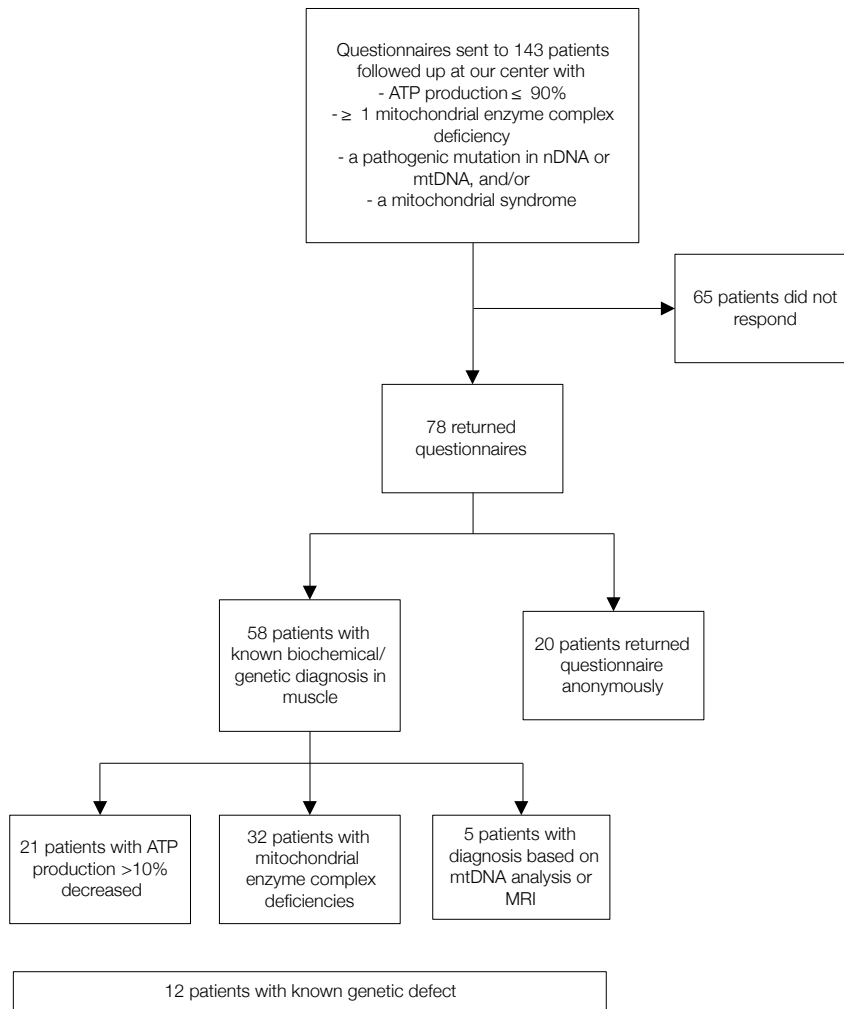
The second questionnaire assessed what impairment was experienced in body functions and activities and participation in daily life. We used the ICF-CY to identify the areas of disability of patients with mitochondrial disorders. The ICF-CY comprises four dimensions: i) body functions; ii) body structure; iii) activities and participation; and iv) environmental factors (Workgroup ICF-CY 2003). We only used the body functions and activities and participation domains. For the activity domains, separate questionnaires for children under 3 years, from 3-6 years, from 7-12 years and from 13 to 18 years were used (Supplementary Document 3). All dimensions include sub domains, which can be assessed using the ICF-CY checklist (e.g. 'Does your son/daughter have any problems being alert and awake?'). We asked the patients or his/her parent(s) to indicate whether a function/activity was affected (0 = not) and, if present, to what extent (1-10 point Likert scale) this was a problem to them. Data of the basic tasks: eating, walking, talking, comprehension, going to school, forming and keeping social relationships and participating in family and community life were explored for differences among groups.

We separately analysed the results of these questionnaires for each age group requiring another ICF-CY list, for the most prevalent clinical syndromes (myopathy, encephalopathy, encephalomyopathy, gastrointestinal involvement and Leigh syndrome). We also separately analysed the results of the first questionnaire in patients rating more than the median amount of symptoms and those rating less than the median amount of symptoms. Burden is defined as the product of prevalence (*i.e.* percentage of people with ICF-CY domain score >0) and the median domain score of patient with non-zero values. The total scores were summed. Since the amount of questions differs per age group, percentages of the maximum score were given as well. Binary values such as the prevalence of the two most common and the two most burdensome symptoms was compared between different groups using the Chi-square test (or Fisher's exact test, depending on the minimum expected cell count). If significant (groups were significantly different from each other) the standardised residuals were interpreted. All (ordinal) ICF-CY domain-parameters were assessed, but all deviated strongly from (log) normality. Therefore, median scores with ranges are given and between-group comparison was performed by the Wilcoxon non-parametric test (2 groups) or the Kruskal-Wallis non-parametric test (>2 groups). In case of significant Kruskal-Wallis ANOVA analysis a posthoc test were performed using the Wilcoxon-Mann Whitney U test with Bonferroni adjustment for the threshold of significance (*i.e.* 0.05/number of comparisons) is used. All data were analysed using PASW 18.0.  $p < 0.05$  was defined as being statistical significant. This study has been conducted in the Netherlands with approval from the regional Medical Research Ethics Committee.

## Results

### Cohort description

The questionnaires were sent to 143 patients and their parents. Seventy-eight (54.5%) of patients and their parents responded positively, of which five patients aged under three years (6.4%), 26 patients aged 3-6 years (33.3%), 32 patients aged 7-12 years (41.0%) and 15 patients were older than 13 years (19.3%). Of the 58 patients of whom gender was noted, 35 (58.6%) were boys. Twenty patients (25.6%) returned the questionnaire anonymously. In the non-anonymous group, consisting of 58 patients (74.4%), biochemical diagnosis was noted. Twenty patients (34.4%) had only decreased ATP production, 11 patients (18.9%) had a Complex I deficiency, three patients (5.2%) had a Complex III deficiency, one patient (1.3%) had a Complex V deficiency and 17 patients (29.3%) had a combined enzyme deficiency. In 12 out of 58 patients (20.7%), the genetic defect causing the mitochondrial dysfunction was found (4 patients with the m.3243A>G, one patient with a combined m.3243A>G and a mutation in the *POLG* gene, one patient with a mutation in the *AGK* gene, and one patient each with m.13781T>C, m.8993T>G, m.6489C>A, m.8363G>A and m.13513>A, mutations and one patient with a combined

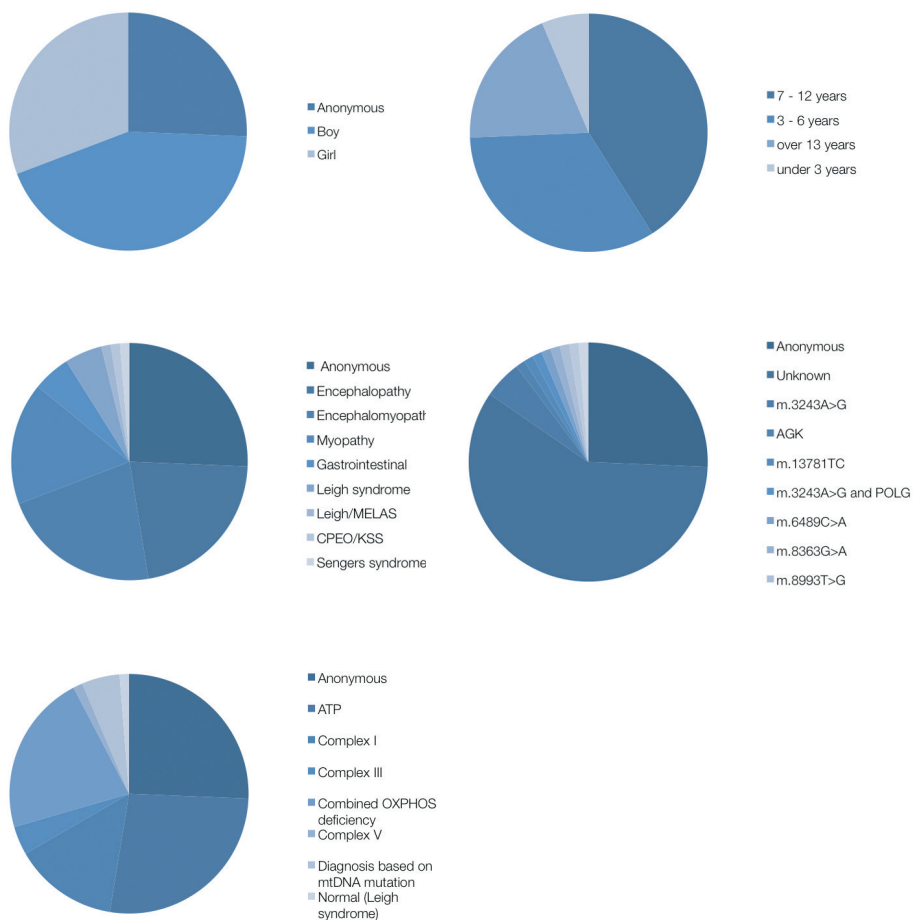


**Figure 3.**

**Flowchart of our study.** 143 patients followed up in our centre were sent two questionnaires, 78 patients responded positively of which 20 patients returned the questionnaire anonymously. Of the remaining 58 patients the muscle biopsy showed: mitochondrial enzyme complex deficiencies in 32 patients, decreased ATP production ( $< 90\%$  of lowest reference value) in 20 patients and in 5 patients, the diagnosis was based on mtDNA or MRI analysis; 12 had a genetic diagnosis.

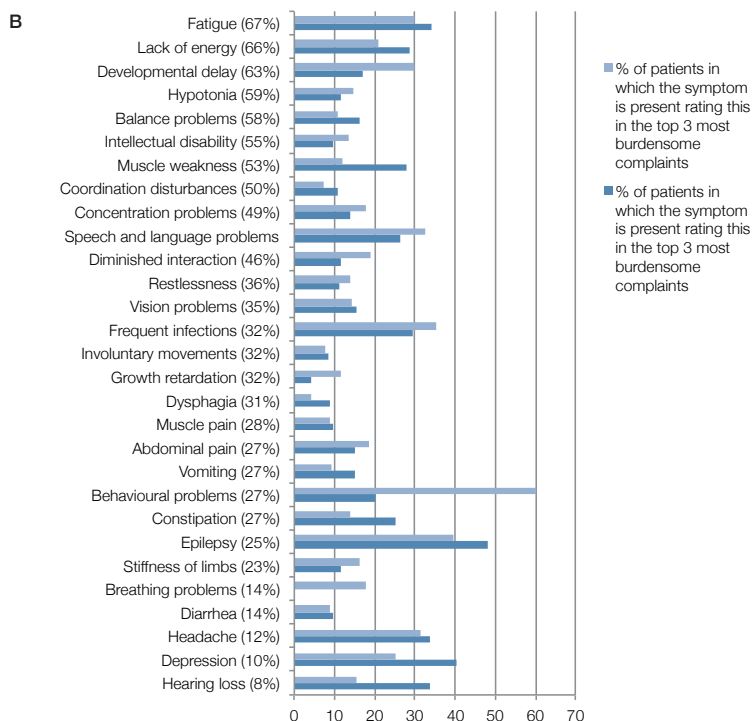
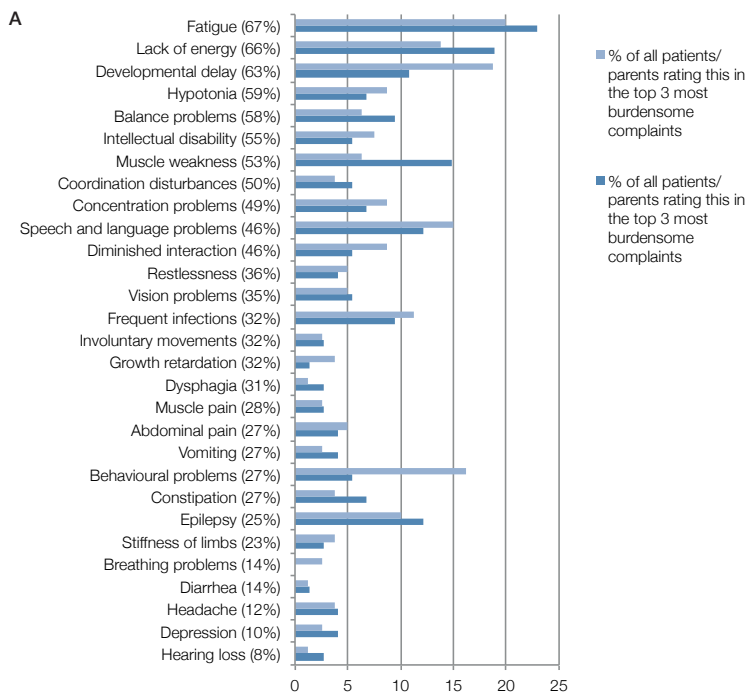


m.3460G>A mutation and 16q23.3 deletion). One patient had a normal biopsy and no mutation in mitochondrial DNA, but did have Leigh syndrome (as evidenced from MRI) with dystonia, epilepsy, tetraplegia therefore fulfilling our criterion iv (Figure 3). For an overview of all patient details (Figure 4).



**Figure 4.**

**Patient description.** A) Gender; B) Age category; C) Biochemical abnormalities in muscle biopsy; D) Genetic defects; and E) Clinical phenotype. AGK = acylglycerol kinase; ATP = adenosine triphosphate; CPEO = chronic progressive external ophthalmoplegia; KSS = Kearns Sayre syndrome; MELAS = mitochondrial encephalopathy, lactic acidosis and stroke-like episodes.



**Figure 5: Presence and burden of symptoms in mitochondrial disease as rated by patients and their parents.** Presence of physical complaints, as reported by parents, in our cohort of patients with mitochondrial disorders, indicated by the percentage behind the complaint (%). The red (upper) bar indicates the percentage of parents rating this symptom in the top 3 of most burdensome symptoms. The blue (lower) bar indicates the percentage of children (> 3 years) rating this symptom amongst the most burdensome symptoms.

## Symptoms

Seventy-eight patients and parents responded to our request to indicate the presence of signs and symptoms of their children. In patients aged 0-3 years, the most prevalent symptoms were developmental delay and balance problems (both 4 out of 5 patients) and constipation, fatigue, hypotonia, muscle weakness, intellectual disability, coordination problems and growth retardation (3 out of 5 patients). In patients aged 3-6 years, the most prevalent symptoms were balance problems (65.4%) and developmental delay, speech and language problems, hypotonia, coordination problems and fatigue and lack of energy (all 57.7%). In patients aged 7-12 years, the most prevalent symptoms were developmental delay (65.6%), fatigue and lack of energy (65.6 % and 68.8% respectively), hypotonia (65.6%), and problems relating to others (59.4%). In patients aged above 13 years, the most prevalent symptoms were lack of energy and fatigue (86.7% and 73.3% respectively), concentration problems, and muscle weakness (both 60%). There was no difference in the prevalence of fatigue, lack of energy, speech and language problems and behavioural problems among the different age groups (Figure 5).

## Symptoms (parents expect) children would like to change

Ninety-one percent (71 out of 78) of the parents and their children filled out the questionnaire concerning what three symptoms and complaints (parents expect) the child would like to change. Six patients and their parents scored only one item, ten patients and their parents filled in two items, and one patient/parent filled in five symptoms. Of the children aged 0-3 years, 1 parent gave no answers and two parents gave all the same answers as their children. Since this indicates the imaginable difficulties estimating the opinion of their young child, we excluded these answers from the analysis. In children aged 3-6 years, the most burdensome symptoms (according to the patient or their parents' estimation) include: speech and language problems (30.8%), fatigue and lack of energy (26.8% and 19.2%, respectively), and developmental delay (26.9%). In children aged 7-12 years, the most burdensome symptoms (according to the patient or their parents' estimation) were fatigue and lack of energy (31.3% and 25%, respectively), decreased muscle power, and epilepsy (both 25%). In children aged over 13 years, the most burdensome symptoms (according to the patient or their parents' estimation) were fatigue and lack of energy (26.7% and 53.3%, respectively), frequent infections, muscle weakness and hypotonia, muscle pain and speech and language problems (all 13.3%; Figure 5A). There was no difference in the burden of fatigue, lack of energy and speech and language problems among children within the different age groups or with more or less than the median amount of symptoms.

## Symptoms parents would like to change

Ninety-five percent (74 out of 78) of the parents filled in the questionnaire about what three symptoms they would like to change. Four parents scored only one item, twelve parents filled in two items, two filled in four items, and one filled in nine items. The latter questionnaire was excluded from the analysis.

Parents of children aged 0-3 years rated muscle weakness (3 out of 5 patients), developmental delay, speech and language problems and frequent infections (all 2 out of 5 patients) as the most important symptoms they would like to change. In children aged 3-6 years, the most burdensome symptoms to parents were speech and language impediments (30.8%), lack of energy (23.1%) and developmental delay and muscle weakness (both 19.2%). In children aged 7-12 years, parents would most like to change behavioural abnormalities (28.1%), fatigue and lack of energy (25% and 21.9% respectively), developmental delay and problems in social interactions (both 15.6%). In children aged over 13 years, the most burdensome symptoms to parents were behavioural problems (33.3%), fatigue and lack of energy (40% and 20%, respectively) and intellectual problems (26.7%; Figure 5B and Table 1). There was no difference in the burden of fatigue, lack of energy, behaviour problems and speech and language problems among parents of children within the different age groups or with more or less than the median amount of symptoms.

## Clinical syndromes

In patients with myopathy ( $n = 13$ ), the most prevalent symptoms were hypotonia, muscle weakness (both 76.9%) and lack of energy and fatigue (both 69.2%). In patients with encephalopathy ( $n = 16$ ), the most common problems include developmental delay (93.8%), intellectual retardation and hypotonia (both 75%), problems in social interactions, and balance and coordination problems (all 68.8%). In children with encephalomyopathy ( $n = 17$ ), the most common symptoms included balance problems (76.5%), fatigue and lack of energy, developmental delay and intellectual problems (all 70.6%), hampered coordination and hypotonia (both 64.7%), concentration problems and speech and language deficiencies (both 58.8%). In children with primary gastrointestinal involvement ( $n = 4$ ), the most common complaints include fatigue and lack of energy and fatigue (both 3 out of 4). The most common symptoms in patients with Leigh or Leigh/MELAS syndrome ( $n = 5$ ) include speech and language deficiencies and difficulties with balance (both 4/5) and tiredness, intellectual disability, coordination problems, stiffness of arms and legs, epilepsy and constipation (all 3/5).

Of the children with myopathy, (their parents estimated that) 46.2% would like to change (to their parents' estimation) their fatigue, 38.5% would like to change muscle weakness and 23.1% would like to change lack of energy, developmental delay and coordination problems. (Parents of) children with encephalopathy (estimated that), 31.1% would like

to change vision problems, 25% would like to change epilepsy and 18.8% would like to change the developmental delay. Parents of children with encephalomyopathy reported that the (estimated) most burdensome symptoms for their children were: muscle weakness (23.5%), vomiting, concentration problems, speech and language impediments, lack of muscle power and energy and epilepsy (all 17.6%). The children with gastrointestinal involvement would (to their parents' estimation) like to change fatigue, lack of energy, concentration problems, depressed mood, and frequent infections (all 1/4). Parents of a child with Leigh syndrome expected that their child would like to his/her change speech abnormalities (2/5), developmental delay, diminished social interaction, intellectual disability, muscle weakness, epilepsy, dysphagia, and breathing abnormalities (all 1/5).

The parents of children with myopathy would like to change fatigue and frequent infections (30.8%), as well as behavioural problems, decreased muscle power, and lack of energy (23.1%). The parents of children with encephalopathy would like to change the developmental delay and epilepsy (both 31.3%) and vision, speech and language abnormalities, as well as behavioural and concentration problems (18.8%). The parents of children with encephalomyopathy found speech and language problems (35.5%), developmental delay (29.4%) and behavioural problems, muscle weakness, and frequent infections (all 23.5%) to be the most burdensome symptoms. The parents of children with primarily gastrointestinal involvement preferred to change fatigue (2/4), lack of energy, frequent infections, social interaction problems, speech and language abnormalities, concentration problems, depressed mood, and headache (all 1/4). The parents of children with Leigh syndrome would like to change speech and language impediments (3/5), and intellectual disability, developmental delay, epilepsy, muscle weakness, involuntary movements, breathing problems and dysphagia (1/5 each). There was no significantly different frequency of fatigue, lack of energy or behaviour or speech problems among the clinical entities. No differences were found in the percentage of children scoring speech problems, fatigue or lack of energy among the most burdensome symptoms. From a parents' perspective, no differences were found in behaviour, fatigue and lack of energy, but a significantly non-equal distribution was seen in speech and language impediments (significantly lower proportion of the parents of a child with myopathy that scored this symptom among the most burdensome symptoms, (standardised residuals -1.7) compared to parents of a child with another phenotype).

### Physicians' perspective

All five paediatricians filled out the questionnaire concerning their views on the three most burdensome symptoms of patients with mitochondrial disorders. Four out of five paediatricians scored fatigue and lack of energy as one of the three the most important items children and parents would like to change, followed by muscle pain and muscle weakness (2 out of 5) in children and developmental delay and muscle weakness (2 out

of 5) in parents (Table 1). One can appreciate the correct estimation of burden in fatigue, lack of energy, developmental delay and muscle weakness. However, there was an overestimation of the burden of muscle cramps and an underestimation of the burden caused by behaviour problems, speech and language problems, epilepsy, hypotonia and frequent infections.

**Table 1.**

**Agreement on most burdensome symptoms among patients and their parents and paediatricians.** Top 10 symptoms rated by parents and patients with prevalence of patients/parents mentioning this symptom in the top-3 most burdensome complaints (% of whole group). Top 7/9 symptoms rated by paediatricians as most burdensome (number of paediatricians rating this symptom). One can observe the correct estimation of burden in fatigue, lack of energy, developmental delay and muscle weakness, the overestimation of the burden of muscle cramps and the underestimation of the burden caused by behaviour problems, speech and language impediments, epilepsy, hypotonia and frequent infections.

Patients (to parents' estimation)	(%)	Patients to paediatricians' estimation	(n =)
Lack of energy	27	Fatigue	4
Fatigue	27	Lack of energy	4
Muscle weakness	18	Muscle pain	2
Speech and language impediments	17	Muscle weakness	2
Developmental delay	14	Speech and language impediments	1
Epilepsy	13	Epilepsy	1
Hypotonia	10	Headache	1
Frequent infections	9		
Concentration problems	9		
Diminished interaction	8		

Parents	(%)	Parents to paediatricians' estimation	(n =)
Fatigue	26	Fatigue	3
Behavioral problems	23	Lack of energy	3
Lack of energy	22	Developmental delay	2
Speech and language impediments	21	Muscle weakness	2
Developmental delay	19	Muscle pain	1
Hypotonia	15	Speech and language problems	1
Muscle weakness	13	Epilepsy	1
Epilepsy	13	Diminished interaction	1
Frequent infections	13	Stiffness of limbs	1
Concentration problems	10		

## Body Function, Activities and Participation

All but two patients (76 patients) filled out the ICF-CY questionnaire about which limitations patients experience in body functioning. Unfortunately, ten parents did not see the daily activities and participation questionnaire on the back of the paper and one only filled it out partially.

In the Body Function domain, the most burdensome complaints included clumsiness and lack of coordination (calculated burden 6.6), body balance and control (6.3), muscles problems (6.3), eating (5.0), and tasks requiring thinking, making sounds and/or saying words (both 4.7). These were also the most prevalent problems. The limitations with the highest median severity score (so if present, most severely affected) include making sounds and saying words (9/10), clumsiness and lack of coordination, bodily balance and control, problems with muscles and eating (all 8/10; Table 2A).

In the Activities and Participation domain, the most burdensome complaints included: learning to write (calculated burden 8.3), learning to calculate (6.9), engaging in activities at school, in the neighbourhood and community (6.8), participating in employment preparedness programs (6.8), walking (6.6) and using the toilet (6.6). The most frequently affected domains include engaging in activities at school, in the neighbourhood and community (85%), learning to write (83%) and dressing one-self (81%). The domains with the highest median score include both learning to write and calculate (both 10/1), learning to read (9.5/10), participating in employment preparedness programs, avoiding self-harm, problem solving (all 9/10), using the toilet and speaking (both 8.5/10) (Table 2B).

**Table 2.**

**Prevalence and severity of impairment.** A) Prevalence of problems in the ICF-FY Body Function domain and B) the Activities and Participation domain. Burden is calculated by multiplying the prevalence and the severity median. In Table 2B the age ranges in which the question is applicable are indicated.

Do you have any problems...	Prevalence of patients with problems in this area (%)	Median score (1-10)	Burden
...with clumsiness or coordinating parts of the body?	82	8	6.6
...with balance or body control?	79	8	6.3
...with muscles of the body, arms or legs?	76	8	6.1
...eating?	62	8	5.0
...with tasks requiring thinking?	67	7	4.7
...making sounds/ saying words?	52	9	4.7
...paying attention to something or someone?	63	6	3.8
Do you have tics, tremors or other unusual movements?	50	7	3.5

...getting the right amount of nutrients?	50	7	3.5
...with bowel movements?	52	6	3.1
...seeing things ?			
...falling asleep or staying asleep?	50	6	3.0
...remembering or recalling something?	48	6	2.9
Do you experience pain more than other children of the same age?	40	7	2.8
...with stiffness of body, arms or legs?	46	6	2.8
...being alert and awake?	41	6	2.5
...with movement of wrists, elbows, shoulders or knees?	49	5	2.5
...controlling arm and leg movements?	48	5	2.4
...with sensitivity or irritation of skin?	42	5.5	2.3
...distinguishing sounds, shapes or smells?	33	7	2.3
...digesting food?	31	6	1.9
Are you allergic or hypersensitive to any food, plant or animal?	24	6	1.4
...hearing sounds ?	29	4.5	1.3
...breathing?	22	5	1.1
...with urination ?	18	5	0.9
...with how the heart functions?	17	4	0.7

Do you have any problems...	Agerange to which question is applicable (years)	Prevalence of patients with problems in this area (%)	Median score (1-10)	Burden
...learning to write?	3-18	83	10	8.3
...learning to calculate?	3-18	69	10	6.9
...in engaging in activities in school, neighborhood or community?	3-18	85	8	6.8
...participating in programs that prepare for employment?	13-18	75	9	6.8
...walking?	0-18	78	8.5	6.6
...using the toilet?	3-18	73	9	6.6
...learning to read?	3-18	69	9.5	6.6
...dressing self?	3-18	81	8	6.5
...participating in the activities of the household?	3-18	77	8	6.2
...washing self?	3-18	77	8	6.2
...relating to others?	0-18	73	8	5.8
...eating?	3-18	72	8	5.8
...avoiding harm to self?	3-18	63	9	5.7
...playing with things?	0-12	70	8	5.6
...playing with others?	0-12	74	7.5	5.6
...speaking?	0-18	65	8.5	5.5



...with concepts such as amount, length, the same or different?	0-12	61	9	5.5
...performing multiple tasks or responding to a command with multiple components?	0-18	68	8	5.4
...participating in (pre-)school education?	3-18	66	8	5.3
...using words, phrases or sentences?	0-12	65	8	5.2
...solving problems?	13-18	56	9	5.0
...sitting up or getting to stand?	0-18	61	7.5	4.6
...being consistent in behavior from day to day?	3-18	57	8	4.6
...with uncontrolled movements of arms or legs?	0-18	65	7	4.6
...using hands and arms?	0-18	59	7	4.1
...performing a single task or responding to a single communication?	0-18	44	8	3.5
...following the requirements of a daily routine?	3-18	46	7.5	3.5
...mouthing, touching or tasting things?	0-12	43	8	3.4
...using gestures, pictures, or drawings to communicate?	0-18	38	8	3.0
...understanding what others say?	0-18	42	7	2.9
...making contact	0-3	50	5	2.5
...understanding the meaning of gestures or pictures?	0-18	35	7	2.5
...making different vocal sounds?	0-18	24	8	1.9

## Body Function, Activities and Participation per age group

In children aged under three years, the most prevalent problems include problems with getting the right amount of nutrients and skin abnormalities (5/5), problems with muscles, balance, eating, playing and sleeping as well as understanding the meaning of concepts such as amount, length, 'the same' and 'different' (4/5). The highest burden was experienced in eating (median 9/10), walking, and proper nutrition and muscles problems (all 8/10), sitting and standing up, clumsiness, and balance problems (all 7/10). In children aged 3-6 years ( $n = 24$ ), the most prevalent symptoms include clumsiness (85%), balance problems (81%), and difficulties in learning to write (77%), read and calculate as well as going to the toilet, dressing oneself, and participating in community activities (all 73%). The most severe limitations experienced were problems with literacy and mathematics (all 8/10), clumsiness (7/10), balance problems, participating in community activities and dressing oneself (all 6.5/10). In children aged 7-12 years ( $n = 32$ ), the most prevalent symptoms include muscle problems (81%), balance problems and clumsiness (both 78%), problems with tasks requiring thinking and participating in household and community activities (all 69%) as well as problems washing and dressing oneself, playing with others, and learning to write (all 66%). The most burdensome domains include difficulties in learning to write (10/10) and mathematics (9/10), washing oneself, and going to the toilet alone, participating in communal activities, understanding concepts

such as amount, length, 'the same' and 'different' and performing multiple tasks and/or responding to multiple comments (all 8/10). In children aged over 13 years ( $n = 15$ ), the most prevalent problems include problems with learning to read and write, problems with muscles, participating in the activities of the community, and forming and keeping social relationships (all 87%) and problems with walking, sitting and standing, balance problems, and limitations in participating in the activities of the family (73%). The most burdensome domains include limitations in participating in employment preparedness programs (9/10), participating in the community activities, and forming and keeping social relationships (both 8/10).

### Body Function, Activities and Participation per clinical diagnosis

In children with myopathy ( $n = 13$ ), the most commonly affected domains included problems with muscles and clumsiness (both 85%), balance problems and limitation in participating in family activities (both 60%). The most burdensome domains included pain, dressing oneself, eating, playing with others and problems performing multiple tasks or responding to a command with multiple components (all 6/10) as well as problems with consistent behaviour (5.5/10). In children with encephalopathy ( $n = 16$ ), the most common limitation include comprehension, clumsiness, and tasks requiring thinking (all 88%), problems with vision, balance and muscles and performing multiple tasks or responding to a command with multiple components (all 81%). The most burdensome domains include performing multiple tasks or responding to a command with multiple components, problem solving, speaking, washing oneself, going to the toilet, preventing harm to oneself, forming and keeping social relationships, going to school and participating in employment preparedness programs (all 10/10). In children with encephalomyopathy ( $n = 17$ ), the most prevalent problems include balance problems (94%), clumsiness (88%), muscle problems (76%) as well as problems with walking and tasks requiring thinking (65%). The most burdensome complaints included limitations in saying words and understanding gestures and learning to write (all 10/10) and problems with walking, talking and eating (9/10). In children with a gastrointestinal phenotype ( $n = 4$ ), the most prevalent problems include problems hearing and sleeping (both 3 out of 4). The most burdensome complaints include problems with attending to school (8/10), making social contact (6/10), sleeping problems (5/10) and problems with eating and proper nutrition (both 4/10). In children with Leigh syndrome ( $n = 5$ ), the most prevalent complaints include tics, tremors or other unusual movements, balance problems (both 5 out of 5), and speech impediments (4 out of 5 patients). The most burdensome complaints include clumsiness, vision problems, pain, breathing problems, skin abnormalities, difficulties learning to write, walk, dress and wash oneself (all 10/10). Analysis of the median score of eating, walking, talking, comprehension, attending school, forming and keeping social relationships and participating in family and community life among age groups and clinical diagnosis, revealed that only 2 parameters reach statistical significance. An unequal distribution of the severity of speech impediments

was observed among the groups, though none of these reached significance in the posthoc test. Surprisingly, a lower prevalence and severity of problems with walking difficulties was observed in the group with myopathy, compared to patients with Leigh syndrome ( $p = 0.007$ ;  $n = 10$  and 3).

### Total score on Body Functions and Activities and Participation

The sum score from the Body Functions domain had a median value of 78 (range 0 – 179). The sum score of the Activities and Participation domain had a median value of 102 (range 0 – 290). The total score had a mean of 176 with a range of 8 to 449. In children with myopathy, the median score for Body Functions was 56 (range 11 – 99), 92 (range 0 – 177) for Activities and Participation and 164 (range 11 – 258) for the total score. In children with encephalopathy, the median score for Body Functions was 111 (range 14 – 179), 200 (range 0 – 290) for Activities and Participation and 295 (range 18 – 430) for the total score. In children with encephalomyopathy, the median score for Body Functions was 75 (range 6 – 167), 107 (range 0 – 282) for Activities and Participation and 191 (range 26 – 449) for the total score. In children with Leigh syndrome, the median score for Body Functions was 88 (range 50 – 136), 114 (range 0 – 179) for Activities and Participation and 167 (range 50 – 295) for the total score. In children with mainly gastrointestinal involvement, the median score for Body Functions was 104 (range 51 – 117), 17 (range 10 – 126) for Activities and Participation and 134 (range 61 – 230) for the total score. No statistical difference in total sub domain scores and total score was observed among the different age groups or clinical groups (Table 3).

**Table 3.**  
**Median percentage of maximum score on the Body Function and Activities and Participation domains among the different clinical groups.**

		Myopathy	Encephalopathy	Encephalomyopathy	Gastro-intestinal	Leigh syndrome
Body Functions	Median (%)	22	43	29	30	34
Activities & Participation	Median (%)	32	71	54	6	44
Total Score	Median (%)	29	55	34	19	32

## Discussion

This is the first study investigating the signs, symptoms and limitations of mitochondrial disorders from both a patient and parent perspective. We found that the most burdensome symptoms (according to parents) children experience include fatigue and lack of energy, epilepsy, muscle weakness, and speech and language impediments. For parents, behavioural problems, speech and language impediments, fatigue, hypotonia and developmental delay are amongst the most burdensome symptoms. From the

perspective of the multi-dimensional ICF-CY, we found a high disease burden in the areas of clumsiness and coordination problems, bodily balance and control, eating, muscle problems, tasks requiring thinking and making sounds or saying words (Body Functions domain) and difficulties in learning to write and calculate, engaging in social activities, participating in employment preparedness programs, walking and using the toilet (Activities and Participation domain). The range in limitations experienced in the Body Function and Activities and Participation domains varies from almost none to 449 (with a maximum score of 550 for that age-group) is extremely high. Children with encephalopathy experienced the most (though not significantly more) limitations in daily life, followed by children with encephalomyopathy and Leigh syndrome.

This study also shows that there is a discrepancy between what paediatricians felt were the most burdensome symptoms for patients, compared to the patients themselves. Although they adequately estimated the burden of fatigue, lack of energy, muscle weakness and developmental delay, the importance of speech impediments, behavioural problems, frequent infections, epilepsy, and diminished interaction between patients and their parents was underestimated, while the importance of muscle cramps and muscle weakness was overestimated. This may suggest that patients and their parents are more hampered by limitations in interaction and daily activities, while physicians are more focused on the physical symptoms of the disease. It is important to take this discrepancy into account when selecting outcome measures for clinical trials or when optimizing patient care. The consequence for the process of selecting outcome measures could for example be that patients and parents should be actively involved in the outcome measure selection process to assure that the domains most important to the patient are covered.

The use of the proxy report in the indication of most severe symptoms rated by the child might have negatively influenced the reliability of our results. First of all, parents are also not fully able to determine the burden of disease of their child. In a recent study in patients with Duchenne muscular dystrophy, parents were able to adequately estimate the physical quality of life but underestimated the quality of life of their sons in three other domains, namely moods and emotions, self-perception and bullying.<sup>16</sup> Secondly, the high percentage of 'children' rating behavioural problems as one of the main things they would like to change, probably indicates the difficulty of parents in separating their opinion from their child's.

This study provides a starting point for the mitochondrial disease field to determine which symptoms and disabilities should be covered in clinical trials. For example, since fatigue and lack of energy are the most common and most burdening symptoms in paediatric mitochondrial disease patients and their parents, it is important to validate or develop adequate outcome measures for these symptoms, such as the Checklist

Individual Strength and the PedsQL multi-dimensional fatigue scale for subjective fatigue<sup>17-22</sup> and the accelerometer to measure daily physical activity.<sup>23,24</sup> Given the high impact of mitochondrial diseases on the activities in daily life and the limited correlation between muscle power and functional abilities,<sup>25</sup> we would advise to use a detailed questionnaire such as the Pediatric Evaluation of Disabilities Inventory (PEDI)<sup>26</sup> for the assessment of patient important treatment effects.

Moreover, the use of outcome measures measuring speech and language development and behavioural problems may measure differences very important to parents. However, the results can also be applied into daily clinical practice, in the form of an intensive rehabilitation trajectory including the consultation of psychologists and speech therapists. In our personal experience, intelligence in children with severe language impairment is frequently underestimated and these children and their parents might benefit significantly from speech devices, for example.

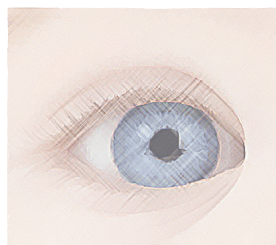
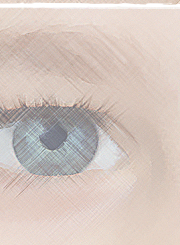
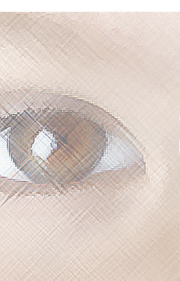
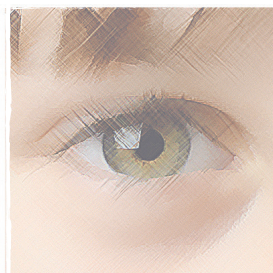
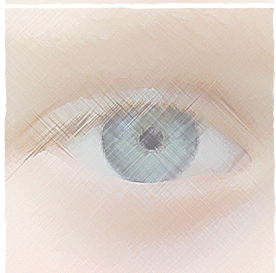
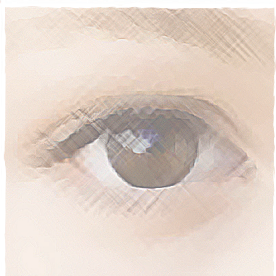
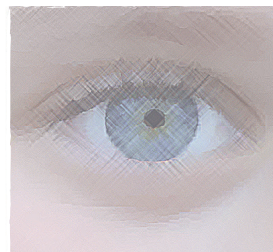
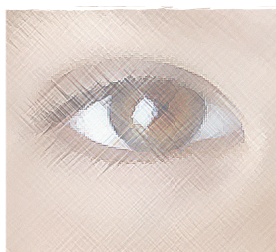
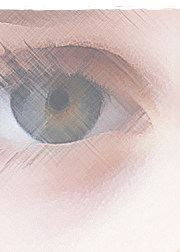
The heterogeneity within our patient cohort allows us to form a good impression of the wide variation of complaints, disabilities and abilities between patients with mitochondrial disorders. From the figures of this study, the challenges and limitations experienced by the wide variety of patients with mitochondrial disorders seem quite homogeneous. However, the results of this study may not be applicable to small homogeneous cohorts of patients. Since clinical phenotypes and cultural differences vary between and within countries, we decided to only include Dutch patients and their paediatricians. We analysed several clinical phenotypes separately, to give an impression of complaints in homogeneous cohorts.

The FDA strongly advises to include outcome measures of importance to patients in every trial. Here, we investigated which symptoms are most burdensome to patients and their parents and which activities and participations in daily life are most restricted. These symptoms include not only the fatigue, lack of energy, muscle weakness and developmental delay, as correctly estimated by physicians, but also a high burden of behavioural, speech and language impediments was reported. This provides not only a starting point for the selection of outcome measures of importance to patients for future clinical trials, but also a significant concern and starting point for supportive therapy when caring of these patients.

## References

1. Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders-past, present and future. *Biochimica et biophysica acta* 2004;1659:115-20.
2. Koene S, Smeitink J. Mitochondrial medicine: entering the era of treatment. *Journal of internal medicine* 2009;265:193-209.
3. Viscomi C, Bottani E, Civiletto G, *et al.* *In vivo* correction of COX deficiency by activation of the AMPK/PGC-1alpha axis. *Cell metabolism* 2011;14:80-90.
4. Wenz T. PGC-1alpha activation as a therapeutic approach in mitochondrial disease. *IUBMB Life* 2009;61:1051-62.
5. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med* 2012;366:1132-41.
6. Hirano M, Garone C, Quinzii CM. CoQ(10) deficiencies and MNGIE: Two treatable mitochondrial disorders. *Biochimica et biophysica acta* 2012;1820:625-31.
7. Koga Y, Povalko N, Nishioka J, Katayama K, Yatsuga S, Matsuishi T. Molecular pathology of MELAS and l-arginine effects. *Biochimica et biophysica acta* 2012;1820:608-14.
8. Chinnery P, Majamaa K, Turnbull D, Thorburn D. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2006:CD004426.
9. Pfeffer G, Majamaa K, Turnbull DM, Thorburn D, Chinnery PF. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2012;4:CD004426.
10. Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. *The Journal of urology* 1998;159:1988-92.
11. Vogelzang NJ, Breitbart W, Cella D, *et al.* Patient, caregiver, and oncologist perceptions of cancer-related fatigue: results of a tripart assessment survey. *The Fatigue Coalition. Seminars in hematology* 1997;34:4-12.
12. Schnadig ID, Fromme EK, Loprinzi CL, *et al.* Patient-physician disagreement regarding performance status is associated with worse survivorship in patients with advanced cancer. *Cancer* 2008;113:2205-14.
13. McFarland R, Taylor RW, Turnbull DM. A neurological perspective on mitochondrial disease. *Lancet Neurol* 2010;9:829-40.
14. Ronen GM, Fayed N, Rosenbaum PL. Outcomes in paediatric neurology: a review of conceptual issues and recommendationsThe 2010 Ronnie Mac Keith Lecture. *Developmental medicine and child neurology* 2011;53:305-12.
15. Rodenburg RJ. Biochemical diagnosis of mitochondrial disorders. *Journal of inherited metabolic disease*.
16. Houwen-van Opstal SL, Jansen M, van Alfen N, de Groot IJ. Health-related quality of life and its relation to disease severity in boys with Duchenne muscular dystrophy: satisfied boys, worrying parents-a case-control study. *Journal of child neurology* 2014;29:1486-95.
17. Varni JW, Burwinkle TM, Szer IS. The PedsQL Multidimensional Fatigue Scale in paediatric rheumatology: reliability and validity. *J Rheumatol* 2004;31:2494-500.
18. Gordijn SM, Cremers EM, Kaspers GJ, Gemke RJ. Fatigue in children: reliability and validity of the Dutch PedsQL(TM) Multidimensional Fatigue Scale. *Qual Life Res* 2011;20:1103-8.
19. Varni JW, Limbers CA, Bryant WP, Wilson DP. The PedsQL Multidimensional Fatigue Scale in type 1 diabetes: feasibility, reliability, and validity. *Paediatric diabetes* 2009;10:321-8.

20. Varni JW, Limbers CA, Bryant WP, Wilson DP. The PedsQL multidimensional fatigue scale in paediatric obesity: feasibility, reliability and validity. *Int J Paediatr Obes* 2010;5:34-42.
21. Swanink CM, Vercoulen JH, Bleijenberg G, Fennis JF, Galama JM, van der Meer JW. Chronic fatigue syndrome: a clinical and laboratory study with a well matched control group. *Journal of internal medicine* 1995;237:499-506.
22. Ketelslegers IA, Catsman-Berrepoets CE, Boon M, *et al.* Fatigue and depression in children with multiple sclerosis and monophasic variants. *Eur J Paediatr Neurol* 2010;14:320-5.
23. Bjornson KF. Physical activity monitoring in children and youths. *Paediatr Phys Ther* 2005;17:37-45.
24. Jeannet PY, Aminian K, Bloetzer C, Najafi B, Paraschiv-Ionescu A. Continuous monitoring and quantification of multiple parameters of daily physical activity in ambulatory Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2011;15:40-7.
25. Merlini L, Bertini E, Minetti C, *et al.* Motor function-muscle strength relationship in spinal muscular atrophy. *Muscle & nerve* 2004;29:548-52.
26. Feldman AB, Haley SM, Coryell J. Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. *Phys Ther* 1990;70:602-10.





# Chapter 3

Towards the harmonisation  
of outcome measures in  
children with mitochondrial  
disorders



## Towards the harmonisation of outcome measures in children with mitochondrial disorders

**Saskia Koene**<sup>1</sup>, Merel Jansen<sup>2</sup>, Chris M. Verhaak<sup>3</sup>, Remco L.A. De Vruet<sup>4</sup>, Imelda J.M. de Groot<sup>1,2</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc

<sup>2</sup>Department of Rehabilitation, Radboudumc

<sup>3</sup>Department of Medical Psychology, Radboudumc

<sup>4</sup>Rare disease matters, Leiden, the Netherlands

Developmental Medicine and Child Neurology 2013 Aug;55(8):698-706

A clinical trial is only as reliable as its outcomes, therefore the careful and systematic selection of outcome measures is extremely important. Currently, the selection of outcome measures for clinical trials designed to evaluate new drugs in patients with mitochondrial disorders is inefficient and has not been addressed systematically. Given that meaningful data can be obtained only from trials in which outcomes are assessed using valid instruments, one should first focus on the validation of a set of selected instruments in the target population. The aim of this review is to systematically select a 'toolbox' of robust outcome measures that are relevant to all patients. Using an extensive search of published literature, we systematically compiled a toolbox with outcome measures based on a primary search for possible instruments. Subsequently, we reduced this toolbox using strict criteria that were adapted from the United States Food and Drug Administration. The result is a toolbox with clinically relevant and psychometrically robust instruments for performing clinical research in children with mitochondrial disorders was compiled. In coming years, more experience using these outcome measures in children with various mitochondrial disease phenotypes must be obtained before reliable conclusions regarding the validity of these instruments can be drawn.

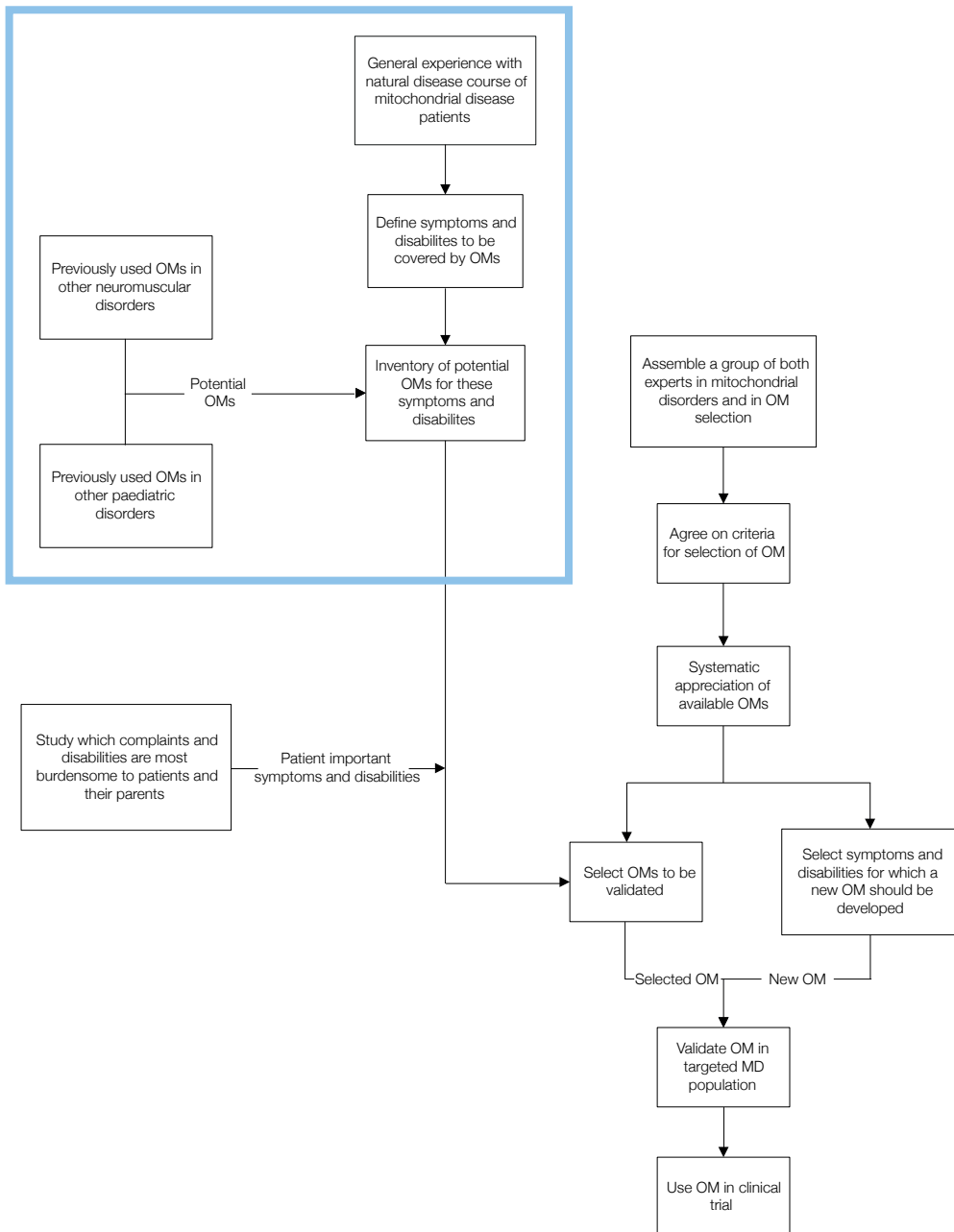
## Introduction

Mitochondrial diseases are the most prevalent inherited metabolic diseases, with an incidence of approximately 1 in 5,000 live births.<sup>1</sup> Most individuals with mitochondrial diseases have a defect in the mitochondrial oxidative phosphorylation (OXPHOS) system, the final biochemical pathway in adenosine triphosphate (ATP) synthesis. Because mitochondria are present in nearly every cell, in principle, symptoms can arise from virtually every organ, although the most commonly affected organs are the brain, heart, and skeletal muscle.<sup>2</sup> In general, two primary phenotypes – mitochondrial encephalopathy and mitochondrial myopathy – can be distinguished, although most children have a combination of both phenotypes, either with or without a multi-system disease. In addition to wide variability in the pattern of the affected organs, the degree of disability also varies widely. While some patients can remain in a mainstream school and achieve normal milestones (e.g. obtaining a swimming certificate), others can barely interact with their environment. Although some rare exceptions exist<sup>3,4</sup>, there is currently no treatment for most mitochondrial diseases.

**Table 1.**  
**Hurdles in outcome measure selection and development in paediatric mitochondrial disease.**

Rare disorders, small groups for validation
Clinically, biochemically and genetically heterogeneous disorders
Clinical heterogeneity within one genetic entity, both in mtDNA and nDNA mutations
Unpredictable, variable and sometimes oscillating natural disease course
Multi-system disease with variable involvement of organs
Patients are often unable to communicate
Patients are often very disabled and unable to perform functional tests (floor effect)
The cooperation required for some tests may be impossible or unreliable in young children, or children with mental retardation or behavioural problems
Abilities of patients vary widely (floor and ceiling effect)
Outcome measure instruments mostly have a limited age range
Children are developing and growing

A recent Cochrane review stated that, although the number of trials has doubled since 2006,<sup>5</sup> the proportion of trials with an adequate trial design has remained extremely low (accounting for only 12 out of 1,335 studies).<sup>6</sup> Assessing the effect of a given treatment in children with a mitochondrial disorder is difficult for many reasons (Table 1). The two major challenges in designing a clinical trial are: i) the wide heterogeneity within this group of disorders, which requires the outcome measures to be almost individually tailored; and ii) the unpredictable, variable, and oscillating clinical disease course, which can require impracticably large cohorts to adequately test the potential benefits



**Figure 1.**

**Flowchart for selecting the outcome measure instruments to be used in clinical trials for children with mitochondrial disorders.** Outcome measures are tests that are used to determine disease severity and to follow the progression of the disease. The position of this study is indicated by the red box. MD: mitochondrial disease; OM: outcome measure.

of a new therapy. A potential solution to these challenges is the careful selection or development of universally applicable and sensitive outcomes and outcome measures (Figure 1).

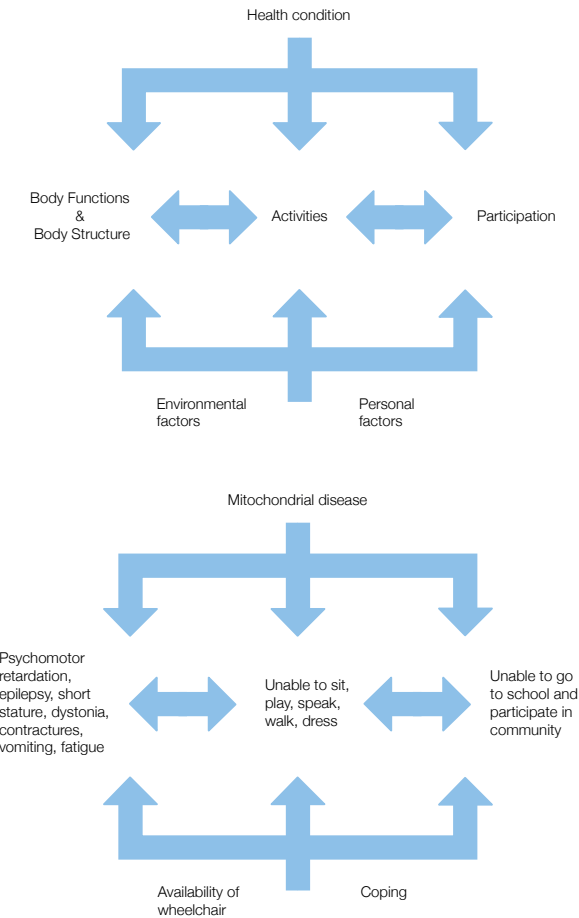
In the field of mitochondrial diseases, a lack of consensus regarding which outcomes measures should be used in clinical trials has resulted in several problems (Table 2), including inefficient trial design and inconclusive trial results. The latter is of major concern to researchers and practitioners, not only because it can lead to the inefficient spending of limited resources, but also because it can be predicted in advance that some trials will yield inconclusive results, which is clearly not consistent with good clinical practice.

**Table 2.**  
**Consequences of the lack of consensus on outcome measures.**

Inefficient, duplication of effort with every trial/clinical study
Suboptimal selection of outcome measures for clinical trials
most test are not clinically relevant to the patient
most tests have not been validated before use in clinical trial
natural fluctuation (due to tiredness, growth and development) of test unknown
no empirical evidence of responsiveness to change, even if valid and reliable
Only scientifically sound tests and trials provide meaningful data
patients shouldn't be burdened by unnecessary tests/trials (Good Clinical Practice) - primum non nocere!
inefficient spending of sparse resources
Only consensus on outcomes will allow future meta-analyses

The aim of this review is to provide a systematically composed set of scientifically robust and clinically significant outcome measures for the entire disease spectrum of paediatric mitochondrial disorders. Because the phenotypes that are associated with mitochondrial disorders are highly heterogeneous, the selected outcome measures should ideally cover the complaints and disabilities that are most prevalent and prominent among children with mitochondrial disorders and should not focus on one specific symptom seen only in a rare mitochondrial syndrome. This generalised approach requires a multi-dimensional framework that shifts the focus from the symptoms, aetiology, and pathophysiology – all of which can vary from patient to patient – to the impact of physical and cognitive limitations experienced in daily life. To achieve this goal, we applied the widely used International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) framework (Figure 2), which was developed by the World Health Organization<sup>7</sup> and is a multi-dimensional framework with a holistic and biopsychosocial vision of the consequences of disease.

As described previously, the applicable domains to be covered were defined with a subsequent review of the available instruments.<sup>8-11</sup> Because no previous study has investigated the robustness of outcome measures in a mitochondrial disease, we used the results of studies of other paediatric neuromuscular and central nervous system disorders.



**Figure 2.**  
**The International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) framework.** The structure of the ICF-CY is shown on the (left). At the right, the framework is completed to show the symptoms, limitations, and disabilities that can be observed in patients with a mitochondrial disease. Note the interplay between the health condition on physical function, activity, and participation and the influence of environmental and personal factors.



## Methods

### Identification of disabilities that can occur in children with a mitochondrial disease

We used the ICF-CY framework (Figure 2) to identify systematically the symptoms and disabilities that occur in the majority of patients who have a mitochondrial disorder. The ICF-CY comprises the following four dimensions: i) body functions; ii) body structure; iii) activities and participation; and iv) environmental and personal factors. Each of these dimensions includes several sub domains (e.g. walking is a sub domain in the activities and participation domain). We assessed the presence of abnormalities that are frequent among patients with a mitochondrial disease in all four domains (and their sub domains) based on a literature search and expert opinions (JS, SK, separate assessments followed by mutual agreement). Because quality of life (QoL) in general was not included as a domain in the ICF-CY, we added health-related quality of life (parent and/or patient reported) to the classification system. General mitochondrial disease severity was also included as a domain.

For all of the domains in this list (Table 3), outcome measure instruments were identified primarily by performing a thorough review of the published literature. In addition, three outcome measures were identified based on the experience of the authors. A flow chart detailing the selection process is presented in Figure 3.

### Outcome measure identification

We searched PubMed for possible outcome measure instruments for each symptom using the following search strings for paediatrics: ((child\*[tiab] OR schoolchild\*[tiab] OR infan\*[tiab] OR adolescen\*[tiab] OR paediatric\*[tiab] OR paediatr\*[tiab] OR boy[tiab] OR boys[tiab] OR boyhood[tiab] OR girl[tiab] OR girls[tiab] OR girlhood[tiab] OR youth[tiab] OR youths[tiab] OR teen[tiab] OR teens [tiab] OR teenager\*[tiab] OR preschool\*[tiab])). In addition, we used the following search strings for outcome measures: (('outcome assessment (health care)'[MeSH Terms] OR (outcome[tiab] AND measur\*[tiab]) OR Evaluation study [tiab] OR 'Clinical Trial' [Publication Type] OR 'trial'[tiab] OR ('rating scale [tiab] or rating [ti])). We used a combination of i) 'the symptom' (including synonyms), ii) the paediatric search string, and ii) the outcome measure search string.

Our initial search yielded more than 24,000 hits; therefore, we chose to review only the first 200 abstracts in each domain to identify outcome measure instruments. For the articles that explicitly focused on outcome measures, the full-text article and its inclusive references were explored, and instruments with good face validity were included (cross-referencing). In addition, outcome measures with good face validity that were listed on the TREAT-NMD Registry of Outcome Measures website (<http://www.researchchrom.com>) were included. If commonly used and valuable instruments were not detected by our search, we included these instruments separately.

<b>Muscle power</b> ("muscle power" or "muscle force" or "muscle weakness")	256	200	10	2003	0	1	11	5	6	1
<b>Muscle tone</b> ("muscle tone" or hypertonia or hypotonia or spasticity or dystonia)	1,095	200	16	2009	1	0	17	10	7	2
<b>Muscle endurance</b> ("muscle endurance" or ("exercise intolerance" and muscle))	24	24	2	1992	0	1	3	2	1	0
<b>Control of voluntary movements</b> ("involuntary movements" or "abnormal movements" or myoclonia or myoclonus or chorea or ataxia)	494	200	7	2008	1	0	8	6	2	0
<b>Changing and maintaining body position</b> (sit or "roll over" or "rolling over" or "motor function")	197	197	5	1985	2	1	8	5	3	4
<b>Walking and moving</b> (walk or walking)	1,166	200	11	2010	0	0	11	9	2	8
<b>Carrying, moving, handling objects</b> ("hand function" or "arm function" or ("handling objects"))	115	115	10	1989	4	0	14	11	3	0
<b>Managing one's own behaviour</b> ("behaviour problems" or "behavior problems" or "beta- vioral problems" or "behavioural problems")	1175	200	13	2009	0	0	13	10	3	0
<b>Self-care</b> ("self-care" or "functional inde- pendence" or (performance and "activities of daily life"))	746	200	7	2008	5	1	13	7	6	1
<b>Participation</b> (participation)	2,179	200	3	2011	0	0	3	0	3	0

Children are developing and growing	Number of results	Number of abstracts reviewed	Different instruments used in PubMed search (n =)	Year of publication of last article reviewed	Added by Research- chROM (n =)	Added by cross referencing (n =)	Total number of instruments	Instruments excluded (n =)	Instruments included (n =)	Instruments included, but in other domain (n =)
<b>Intellectual function</b> (intelligence or "intellectual ability" or IQ or "intelligence quotient")	3,168	200	12	2011	0	1	13	11	2	2
<b>Energy and drive</b> (fatigue or tiredness or "lack of energy" or "daily physical activity")	706	200	14	2010	5	1	20	14	6	0
<b>Psychomotor development</b> (development and (motor or psychomotor))	1,876	200	8	2011	0	1	9	6	3	0
<b>Affective functions</b> (depression or "affective disorder")	4,487	200	9	2011	0	0	9	6	3	0
<b>Speech and language</b> (language delay" or "language problems" or "speech delay" or "speech problems")	138	138	12	1980	0	1	13	10	3	0
<b>Heart contraction force</b> (heart or myocard) and (contractility or "contraction force"))	68	68	7	1976	0	1	8	6	2	1
<b>Exercise tolerance</b> (exercise test" or "exercise tolerance" or endurance or "exercise intolerance")	1,327	200	10	2009	0	0	10	8	2	3
<b>Insufficient energy production</b> (ATP production" or "energy production" or "ATP synthesis")	2	2	0	1997	0	2	2	0	2	0
<b>Accumulation of metabolites in the body</b> (metabolite or metabolites) and accumulation) or (accumulation and mitochondrial dysfunction"))	0	0	0	-	0	2	3	0	3	0

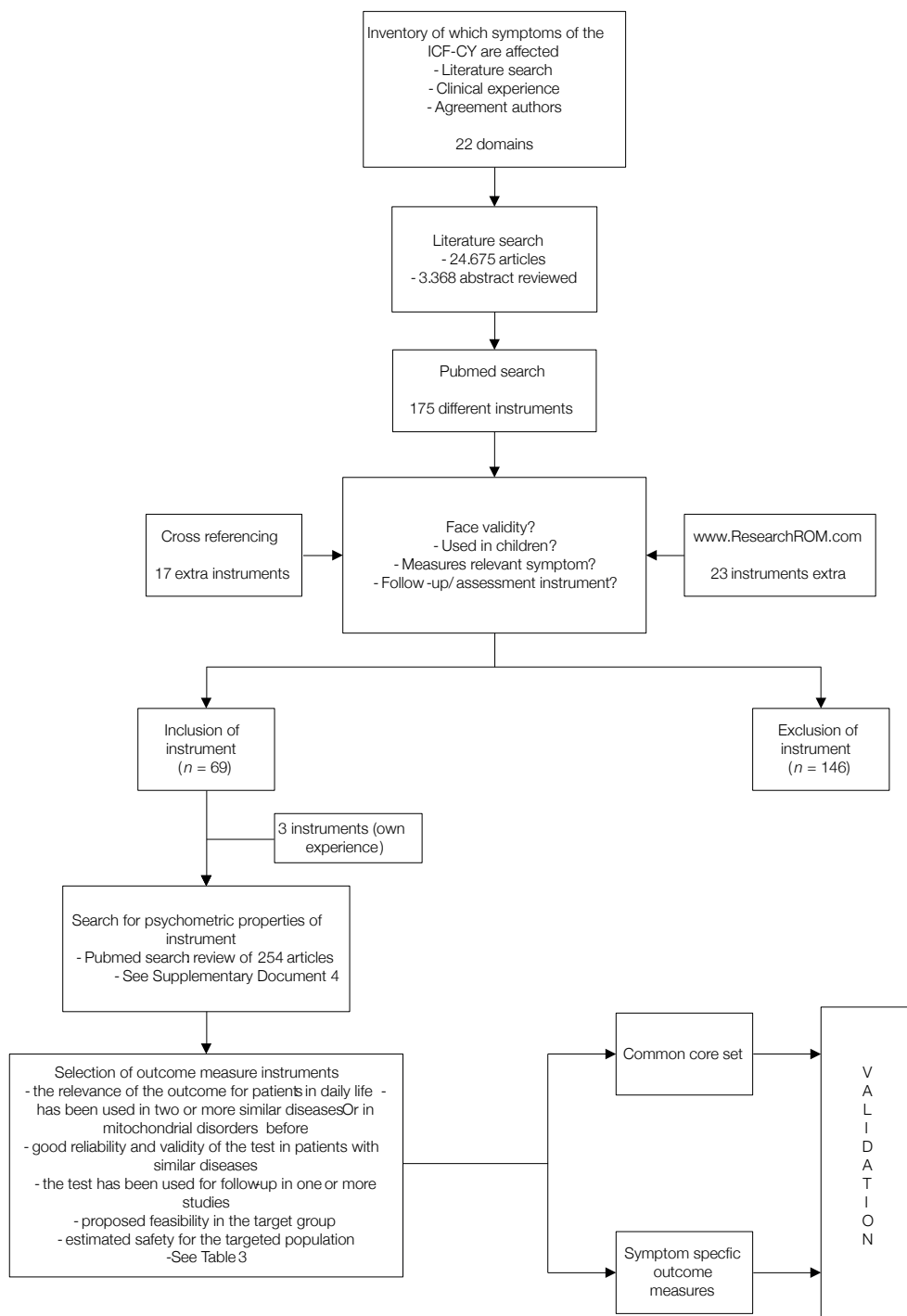
**Table 3.**

**Search strings and number of results.** The selected symptoms/chapters from the ICF-CY are listed in bold, and the search string that was used is shown below each symptom/ chapter. The number of results obtained from PubMed using this symptom/domain search string in combination with the paediatric and outcome string is indicated at the right. Because only the first 200 abstracts were reviewed for each symptom, we list the year of the oldest article that was reviewed. If a scale had multiple scales for different age groups, only one scale was counted. The number of instruments that were found in addition to those identified by the PubMed search, and the number of additional instruments found on researchrom.com are listed; finally, the total number of instruments and the number of excluded and included instruments are indicated. Note: quality of life questionnaires specifically designed for an irrelevant disease (e.g. respiratory disease) are not included in this number.

**Figure 3.**

**Flowchart depicting how the toolbox was created.** First, we inventoried which symptoms/limitations are commonly present in patients with a mitochondrial disorder. Next, the available instruments for these symptoms in children were identified based on a search of the published literature, cross-referencing, the research Registry of Outcome Measures database, and our own experience. We then assessed whether each instrument could be useful in a paediatric mitochondrial disease (face validity). Based on face validity, the instruments were then included or excluded from our review. From our own experience, three tests were added (magnetic resonance imaging, growth charts, and fibroblast growth factor 21). Finally, the 72 instruments that were selected for our review were extensively studied for their psychometric properties (Supplementary Document 3), from which we selected a core set of symptom-specific instruments for clinical trials in children with a mitochondrial disorder. ICF-CY, International Classification of Functioning, Disability and Health for Children and Youth.

<b>Parental/family strain/stress</b> (parental or parenting) and (stress or strain)	1,087	200	9	2010	3	1	13	12	1	0
<b>Quality of life</b> (“quality of life”)	4,345	200	9#	2012	2	2	13	8	5	1
<b>General mitochondrial tests</b> (“mitochondrial disease” or “mitochondrial disorder”)	24	24	1	2003	0	0	1	0	1	1
<b>Total number of articles</b>	<b>24,675</b>	<b>3,368</b>	<b>175</b>		<b>23</b>	<b>16</b>	<b>214</b>	<b>146</b>	<b>68</b>	<b>24</b>



## Outcome measure selection for inclusion in this review

From all of the outcome measures that our search yielded, only the outcome measures that had face validity were included in the review. Face validity was determined using the following criteria: i) previously used in children, and ii) measures symptoms that might be present in any phenotype of mitochondrial disorders, and iii) follow-up/assessment instrument, not a diagnostic instrument or screening instrument for the general population.

## Psychometric data collection

Very few outcome measure studies have been performed in patients with mitochondrial disorders, and consequently the psychometric properties of most tests that are given in this specific patient population are unknown. Therefore, we collected the psychometric data of validation studies in other, better-studied paediatric disorders. This search was performed using the paediatric search strings as described above combined with the name (including the abbreviation) of the test. Where applicable, we included data from similar disorders, including neuromuscular disorders such as Duchenne muscular dystrophy (DMD), spinal muscular atrophy (SMA), and Friedreich's ataxia (FRDA), as well as encephalopathies such as cerebral palsy (CP), as these disorders include symptoms that are similar to mitochondrial disorders such as muscle weakness, coordination problems, spasticity or hypotonia, and a progressive disease course (with the exception of CP, although this disorder was still included, as it is the only common and well-studied central nervous system disorder with symptoms that are common to mitochondrial disorders). This approach probably provided the best indication of the methodological quality and robustness of the test in our population. If the data for these disorders were not available, data from other paediatric conditions were used. A total of 254 articles (Figure 2) were reviewed and yielded sufficient psychometric data for all instruments (Supplementary Document 4).

## Outcome measure appreciation

Following the review of psychometric data, the outcome measures for our toolbox were selected. Because we wanted our toolbox to contain only the most scientifically robust instruments, the instruments were required to fulfil all of our established criteria.

Given that these criteria are quite strict and because we wanted to retain our holistic and multi-dimensional approach, we also provided an additional set of tests containing the outcome measures that fulfilled four or five of our six criteria. To distinguish the instruments based on whether they are robust or not, we used regular and italic typefaces. Therefore, each of the six following criteria (adapted from United States Food and Drug Administration<sup>12</sup>) should be satisfied: i) the outcome is relevant in daily life to a significant proportion of the children with a mitochondrial disorder; ii) the test has been used in mitochondrial disorders or at least one disease with similar symptoms (e.g. DMD, SMA, FRDA, or CP); iii) there is high reliability and validity (inter- and intrarater reliability

and test–retest reliability  $>0.6$  and correlation with other valid instrument measuring the same symptom  $>0.5$ ) in patients with diseases with similar symptoms (e.g. DMD, SMA, FRDA, or CP); iv) the test has been used for follow-up in one or more studies (and has preferably been found to be sensitive to change); v) there is proposed feasibility in the target group; and vi) there is estimated safety for use in the target population.

## Results

### Agreement regarding disabilities in patients with mitochondrial disorders

Two authors (SK and JS, physicians who are involved in the care of patients with mitochondrial disorders) agreed to include the following ICF-CY sub domains: intellectual function, energy and drive, psychomotor development, affective functions, heart contraction force, pulmonary function, insufficient energy production, accumulation of metabolites in the body, muscle power, muscle tone, muscle endurance, and control of voluntary movements (Domain I); structure of the central nervous system (Domain II); changing and maintaining body position, walking and moving and carrying, moving and handling objects, managing one's own behaviour and self-care (Domain III); and parental stress (Domain IV). Additionally, QoL and general mitochondrial outcome scoring lists were included.

### Identification and selection based on face validity

For each domain, we searched PubMed for outcome measures in children. For each symptom, we found 2 to 4,487 articles. Because we found a total of 24,675 articles, we reviewed only the first 200 abstracts for each domain. We therefore reviewed the abstracts of 3,368 (14%) articles and detected 175 outcome measure instruments. Based on the aforementioned criteria for face validity, 68 instruments were selected for a thorough review of their psychometric properties. Twenty-three additional instruments were extracted from the TREAT-NMD Registry of Outcome Measures website database (<http://www.researchchrom.com>) based on face validity, and 16 additional instruments with good face validity were detected using cross-referencing. Three additional instruments were included based on our own experience (fibroblast growth factor 21, magnetic resonance imaging (MRI), and growth charts; Figure 3; Table 2).

### Psychometric properties of instruments

All 72 detected instruments with face validity were then explored individually for psychometric properties in children with or without disease. For some tests, extremely few psychometric data were available, whereas other tests were extensively validated. Only two of the 72 instruments, the Newcastle Paediatric Mitochondrial Disease Scale and the biomarker fibroblast growth factor 21, have been validated in patients with mitochondrial disorders, albeit in only one study each. We summarise all of our results,

including reliability and validity, method of administration, time to administer, age reference, responsiveness, and strengths and weaknesses, in the Supplementary Document 4.

Outcome measure toolbox for children with mitochondrial disorders

Table 4 shows the core set of outcome measures that is recommended for all children with a mitochondrial disorder as well as for children with a (predominant) myopathy (e.g. Kearns–Sayre syndrome) or encephalopathy (e.g. Leigh or mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS)). In our experience, most children have a mixed phenotype; therefore, we recommend determining whether myopathy or pyramidal/extrapyramidal symptoms are the most disabling conditions and then selecting tests based on the abnormalities that are most prominent during the physical examination. We selected outcome measures for the common core set at the ICF-CY subdimensions energy and drive (e.g. accelerometer), psychomotor development (the Alberta Infant Motor Scale), neuromuscular function (e.g. the Medical Research Council Scale), cardiovascular and respiratory system (e.g. speckling tracking MRI), walking and moving (e.g. the 6-minute walking test), and changing and maintaining body position (e.g. the Motor Function Measure). In addition, we selected the KIDSCREEN-52 measurement to assess quality of life. Optional tests for specific groups of children were also selected from the various ICF-CY subdimensions and included a measure of general mitochondrial disease severity (the Newcastle Paediatric Mitochondrial Disease Scale) and a measure of growth state (growth charts). The recommended core set is based on the psychometric properties of the outcome measures (Supplementary Document 4), and the strict selection criteria that are described in the Methods section.

**Table 4.**  
**Suggested outcome measures for children with a mitochondrial disorder.** Common core set (top), as well as optional tests for specific groups and purposes (bottom.). The instruments shown in *italics* met only four or five of the six selection criteria but were included to retain the holistic and multi-dimensional vision on the disease. Note: none of these instruments has been validated for follow-up in children with a mitochondrial disease. Election was based on relevance and feasibility in this patient group and strict criteria for psychometric robustness.



INSTRUMENT	MEASURES	ENCEPHALOPATHY		MYOPATHY		AGE RANGE (YEARS)	TEST TYPE	TIME (MIN)
		MILD	SEVERE	MILD	SEVERE			
KIDSCREEN-52	Quality of Life	+	++	+	+	8 - 18	Q	15-20
Paediatric Evaluation of Disabilities Inventory	Performance and Capability of Activities of Daily Life	+	+	+	+	0.5 - 7.5**	Q	45
PedsQL Multidimensional Fatigue Scale	Subjective fatigue	+	++	++	+	2 - 18	Q	4-5
Accelerometer	Daily physical activity	+	++	+	+	ND	A	
Graded and timed function tests	Motor abilities	+		+	+	> 3	P	5
6 minute walking test	Walking capacity	+		+		> 4	P	10
Gross Motor Function Measure	Gross and fine motor function	+	+			> 0.5	P	45-60
Alberta Infant Motor Scale	Psychomotor development	+	+	+	+	0 - 1.5	O	20-30
Motor Function Measure	Gross and fine motor function		+	+	+	2 - 60	P	15-45
Maximal voluntary isometric contraction	Muscle power, only in MFC≥4			+		> 4	P	5
Medical Research Council Scale	Muscle power, only if MFC<4			+	+	ND	P/O	1
Speckle tracking MRI or 2D/3D strain imaging	Cardiac contractility	+	+	+	+	> 0	I	
Modified Tardieu Scale	Spasticity	+	+			ND	P	5
Movement Disorder Childhood Rating Scale	Movement disorders	+	+			0 - 18	O/Q	10
Jelssen-Taylor Hand function test	Hand function	+	+	+	+	>3	P	30
Forced Vital Capacity	Respiratory muscle power			+	?	> 5	P	10
Sniff Nasal Inspiratory Tests	Respiratory muscle power			+	?	> 4	P	10
Isokinetic dynameter	Muscle power/spasticity	+		+	+	> 4	P	10-15

## COMMON CORE SET

## OPTIONAL TESTS

OPTIONAL TESTS										
Infant Toddler Quality of Life Questionnaire	Quality of Life	+	+	+	+	+	+	2 m -5	Q	5-25
Child Health Questionnaire	Quality of Life	+	+	+	+	+	+	5-18	Q	5-25
Family Strain Questionnaire	Family/parental stress	+	+	+	+	+	+	ND	Q	20
Assessment of Preschool Children's Participation	Participation	+	+	+	+	+	+	2-6	Q	30-40
Wechsler Intelligence Scale for Children IV	Intelligence	+	+	+	+	+	+	6-16	T	65-80
Children's Depression Rating Scale - Revised	Depression	+	+	+	+	+	+	6-12	Q/O	15-20
Reynell Developmental Language Scale III or Test of Early Communication and Emerging Language	Speech and language	+	+	+	+	+	+	1-7	T	25-40
Eyberg Child Behaviour Inventory	Behavioural problems	+	+	+	+	+	+	2-25	Q	5-10
Checklist Individual Strength	Fatigue	+	+	+	+	+	+	> 10	Q	4-5
OMNI scale	Fatigue (during exercise)	+	+	+	+	+	+	8-18	Q	1
Maximal Exercise Test	Exercise capacity	+	+	+	+	+	+	> 4	P	10-30
Assisted 6 minute cycling test for arms and legs	Exercise capacity	+	+	+	+	+	+	4-18	P	30
Newcastle Pediatric Mitochondrial Disease Scale	Mitochondrial disease severity	+	+	+	+	+	+	0-18	P/Q	25
Growth Charts	Growth and nutritional state	+	+	+	+	+	+	0-21	P	5
International Cooperative Ataxia Rating Scale	Ataxia	+	+	+	+	+	+		P	30

2D = two-dimensional; 3D = three-dimensional; a = Only parent report; b = Motor developmental age; A = apparatus; I = imaging/imaging processing; MRC = Medical Research Council; MRI = magnetic resonance imaging; ND = not described/determined; O = Observation; PedsQ = Pediatric Quality of Life Inventory; Q= questionnaire; P = physical test/examination; T = test.

As shown in Table 4, age and disease severity should be taken into consideration when selecting an appropriate instrument as each instrument is applicable only to a specific age and/or disease severity range. Thus, some instruments – including questionnaires that measure QoL, performance and capacity of the activities of daily life and subjective fatigue – can be applied to all patients within the age group for that instrument. For other instruments, however, inclusion or exclusion should be based on the abilities of the patient. For example, for testing ambulatory children with relatively mild myopathy, robust instruments that measure exercise tolerance (e.g. maximum exercise tests or the 6-min walking test) may be feasible. However, more severely affected (wheelchair-dependent) children are unable to perform these tests. It is therefore worth developing or validating new outcome measures such as the three-dimensional accelerometer and the assisted 6-minute cycling test to assess exercise tolerance. Exercise tolerance is a domain that appears to be sensitive to small but relevant treatment effects. Another example is the selection of either the Gross Motor Function Measure (GMFM) or the Motor Function Measure (MFM), based on the presence of spastic or flaccid paresis respectively. Because cardiomyopathy is a life-limiting condition,<sup>12</sup> and because altered myocardial function is present in many mitochondrial disorder patients,<sup>13–15</sup> it seems reasonable to measure myocardial function during clinical trials using instruments that are sensitive to small changes in myocardial contractility, even though the clinical consequences of small changes in myocardial function are currently unknown. Although age at death is commonly used as an outcome measure for lethal diseases, the wide variability in homogeneous biochemical and genetic entities such as Complex I deficiency<sup>16</sup> means that studies should include extremely large groups of patients; however, such large groups are not currently available. Because function does not always follow (abnormalities in) brain structure, the quantification of brain lesions based on MRI was not included in the core set. Moreover, because the spontaneous disappearance of lactate peaks on <sup>1</sup>H-magnetic resonance spectroscopy has been described in progressive Leigh disease,<sup>17</sup> this test was not appropriate for judging treatment effects.

## Discussion

This is the first study to systematically compose a toolbox containing robust and clinically relevant outcome measures that cover virtually the entire disease spectrum of paediatric mitochondrial disorders. Assembling such a toolbox is challenging for mitochondrial disorders, particularly given the heterogeneity and unpredictability of these diseases. We therefore used the multi-dimensional ICF-CY framework to select the most commonly encountered disabilities. We then selected 14 general and 19 symptom specific psychometrically sound and clinically relevant outcome measures by thoroughly reviewing the published literature. The results of this study will facilitate the efficient and universal selection of outcomes for use in future clinical trials, which will in turn improve the quality of future clinical trials involving children with mitochondrial disorders.

It should be noted that the instruments contained within our toolbox have not yet been validated for mitochondrial diseases. However, the validation of a new instrument is complicated by the lack of a criterion standard and the difficulties in defining homogeneous and representative cohorts. Moreover, growth, development, and variable cooperation require age-appropriate instruments and reference values. Therefore, additional experience using these instruments in this population is needed before a definitive toolbox can be created. After sufficient experience has been obtained, and using the multi-disciplinary expertise of outcome measure experts, statisticians and legal authorities, the next steps towards valid outcomes can be taken.<sup>11,18,19</sup>

The instruments in the toolbox were selected based on their applicability to the widest possible range of disabilities. Although this approach has major advantages for making comparisons between groups, the use of a generic instrument can cause reduced specificity, a floor and/or ceiling effect, and decreased sensitivity to small changes that may be clinically relevant to an individual patient. Moreover, many instruments are applicable to patients within only a small age range, and few valid instruments for infants and toddlers exist. Therefore, it may be necessary to include highly sensitive quantifiable methods for measuring daily physical activity in severely disabled patients and exercise tests or muscle power tests in patients with more capabilities, as well as instrument sets for testing both younger and older children.

For all subjective tests, and particularly for questionnaires that assess general well-being, one should consider that the supportive care and optimism associated with participating in a clinical trial can increase the participant's well-being. One should therefore use caution when using these questionnaires in open-label studies. The reliability of a questionnaire in assessing the health-related quality of life for children with severe intellectual disability is unknown; however, in children with reasonable language skills and cognitive development (*i.e.* a developmental age of approximately 5 years), self-

report QoL questionnaires should also be used, as they can provide another perspective of the impact of the disease than parent questionnaires.

In our opinion, the feasibility of using the entire toolbox is quite good. Using the entire toolbox for a patient with myopathy will take approximately 2.5 hours, including more than 1 hour to complete the questionnaires. Because most patients will find these tests physically challenging, we recommend performing the tests in the same sequence, at the same time of day, and on the same day of the week while alternating between physical tests and questionnaires. Most of the tests in the toolbox can be performed by a physiotherapist. Finally, based on the capacities, limitations, and/or other complaints specific to the individual child, certain tests in the toolbox can be either skipped or included. In composing this toolbox, we used a few practical considerations that may influence the applicability of the results. First, we did not seek to provide outcome measures for all symptoms that may be present in the full disease spectrum of mitochondrial disorders; rather, we focused on the symptoms and limitations that are common to most patients. Moreover, the ICF-CY domains were selected based on a physician's opinion, which can differ from the opinions of patients, policymakers, and certain specialists.

Additional studies (e.g. using special focus groups) may reveal discrepancies between these groups with respect to their experiences and interests. In addition, although more than 24,000 articles were detected by our initial search, only the first 200 abstracts for each domain were reviewed; as a result, some instruments that are rarely used may have been missed. Thus, because of the lack of experience in mitochondrial disorders, it is not possible to determine the validity of these instruments in this population using previous studies in which disorders with similar symptoms were used to estimate the robustness of the instruments. However, it should be noted that the validity of a test used in patients with DMD or CP does not necessarily reflect the validity of the same test in mitochondrial disorders, as the phenotypes do not overlap completely, and mitochondrial disorders are often more complex in terms of multi-system involvement. Finally, the multi-dimensionality of the ICF-CY is not reflected in the common core set of our toolbox, as the tests (e.g. for participation or family stress) did not fulfil our extremely strict selection criteria. However, although imperfect, the approach used here will facilitate the selection of outcome measures for rare mitochondrial symptoms and other diseases.

Although many clinical trials have been performed in patients with mitochondrial disorders, most of these trials used an inadequate trial design and/or yielded a non-conclusive outcome.<sup>5,6</sup> Because a growing number of compounds are beneficial in cell and animal models,<sup>2,20,21</sup> we focused on selecting outcomes for clinical trials in humans.<sup>13-15</sup> However, the systematic selection of a toolbox containing clinically relevant and psychometrically

robust instruments is just the beginning, and this study may serve as the starting point for further efforts to obtain international consensus regarding the outcome measures. International collaboration is necessary in order to gain insight into the feasibility and validity of these instruments in well-defined groups of patients with mitochondrial disorders. Only after achieving this goal will clinical trial design be more efficient, and trials will harvest valid data that can ultimately be used in meta-analyses, thereby preventing unnecessary expenses and, most importantly, sparing patients from undergoing unnecessary research procedures.<sup>2</sup>

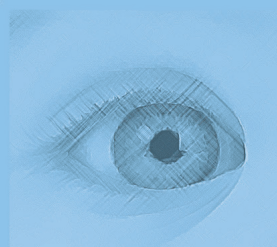
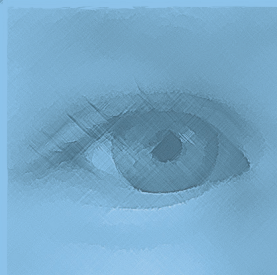
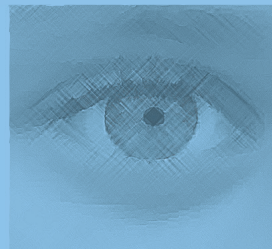
## References

1. Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders – past, present and future. *Biochim Biophys Acta* 2004; 1659: 115–20.
2. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med* 2012; 366: 1132–41.
3. Hirano M, Garone C, Quinzii CM. CoQ(10) deficiencies and MNGIE: two treatable mitochondrial disorders. *Biochim Biophys Acta* 2012; 1820: 625–31.
4. Koga Y, Povalko N, Nishioka J, Katayama K, Yatsuga S, Matsuishi T. Molecular pathology of MELAS and l-arginine effects. *Biochim Biophys Acta* 2012; 1820: 608–14.
5. Chinnery P, Majamaa K, Turnbull D, Thorburn D. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2006; 1: CD004426.
6. Pfeffer G, Majamaa K, Turnbull DM, Thorburn D, Chinnery PF. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev* 2012; 4: CD004426.
7. World Health Organization. ICF: International Classification of Functioning, Disability and Health. Geneva: World Health Organization, 2001.
8. Ronen GM, Fayed N, Rosenbaum PL. Outcomes in paediatric neurology: a review of conceptual issues and recommendations. The 2010 Ronnie Mac Keith Lecture. *Dev Med Child Neurol* 2011; 53: 305–12.
9. Merkel PA, Aydin SZ, Boers M, *et al.* The OMERACT core set of outcome measures for use in clinical trials of ANCA-associated vasculitis. *J Rheumatol* 2011; 38: 1480–6.
10. de Benedictis FM, Guidi R, Carraro S, Baraldi E. Endpoints in respiratory diseases. *Eur J Clin Pharmacol* 2011; 67(Suppl. 1): 49–59.
11. Mercuri E, Mayhew A, Muntoni F, *et al.* Towards harmonization of outcome measures for DMD and SMA within TREAT-NMD; *Neuromuscul Disord* 2008; 18: 894–903.
12. Food and Drug Administration. FDA Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Rockville: Food and Drug Administration, 2009.
13. Scaglia F, Towbin JA, Craigen WJ, *et al.* Clinical spectrum, morbidity, and mortality in 113 paediatric patients with mitochondrial disease. *Paediatrics* 2004; 114: 925–31.
14. Marcus KA, Barends M, Morava-Kozicz E, Feuth T, de Korte CL, Kapusta L. Early detection of myocardial dysfunction in children with mitochondrial disease: an ultrasound and two-dimensional strain echocardiography study. *Mitochondrion* 2011; 11: 405–12.
15. Hollingsworth KG, Gorman GS, Trenell MI, *et al.* Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load. *Neuromuscul Disord* 2012; 22: 592–6.
16. Koene S, Rodenburg RJ, van der Knapp MS, *et al.* Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: what we learned from 130 cases. *J Inherit Metab Dis* 2012; 35: 737–47.
17. Lee SK, Kim J, Kim HD, Lee JS, Lee YM. Initial experiences with proton MR spectroscopy in treatment monitoring of mitochondrial encephalopathy. *Yonsei Med J* 2010; 51: 672–5.

18. European Medicines Agency. (2012) Workshop on Endpoints for Cystic Fibrosis Clinical Trials. [http://www.ema.europa.eu/ema/index.jsp?curl=pages/news\\_and\\_events/events/2012/07/event\\_detail\\_000609.jsp&mid=WC0b01ac058004d5c3](http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/events/2012/07/event_detail_000609.jsp&mid=WC0b01ac058004d5c3) (accessed 11 December 2012).
19. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction. *J Rheumatol* 1993; 20: 528–30.
20. Viscomi C, Bottani E, Civiletto G, *et al.* In vivo correction of COX deficiency by activation of the AMPK/PGC-1 $\alpha$  axis. *Cell Metab* 2011; 14: 80–90.
21. Wenz T, Diaz F, Spiegelman BM, Moraes CT. Activation of the PPAR/PGC-1 $\alpha$  pathway prevents a bioenergetic deficit and effectively improves a mitochondrial myopathy phenotype. *Cell Metab* 2008; 8:249–56.







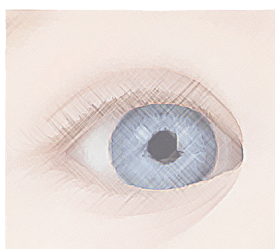
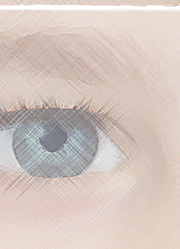
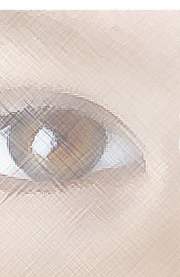
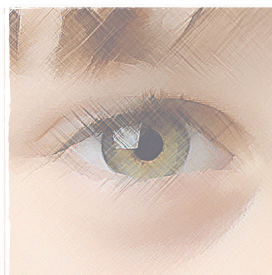
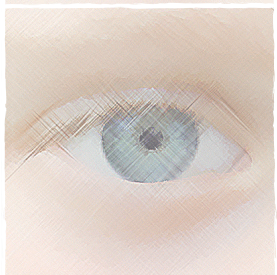
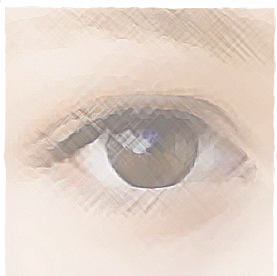
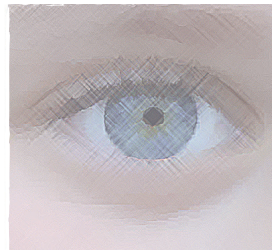
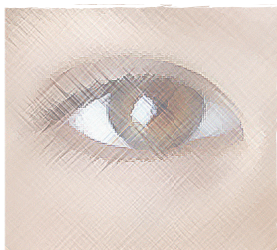
# Part II

## Natural disease course



## Goal

To select inclusion criteria for the validation studies based on the clinical heterogeneity and natural disease course in two different cohorts of patients with mitochondrial disease.



# Chapter 4

Natural disease course and  
genotype-phenotype correlations in Complex I deficiency  
caused by nuclear gene  
defects: what we learned  
from 130 cases





## Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: what we learned from 130 cases

**Saskia Koene**<sup>1</sup>, Richard J. Rodenburg<sup>1</sup>, Marjo S. van der Knaap<sup>2</sup>, Michel A.A.P. Willemsen<sup>3</sup>, Wolfgang Sperl<sup>4</sup>, Vincent Laugel<sup>5</sup>, Elsebet Ostergaard<sup>6</sup>, Marc Tarnopolsky<sup>7</sup>, M.A. Martin<sup>8</sup>, Victoria Nesbitt<sup>9</sup>, Janice Fletcher<sup>10</sup>, Simon Edvardson<sup>11</sup>, Vincent Procaccio<sup>12</sup>, Abdel Slama<sup>13</sup>, Lambert P.W.J. van den Heuvel<sup>1</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc, Netherlands

<sup>2</sup>VU University Medical Centre, Department of Paediatrics, Amsterdam, the Netherlands

<sup>3</sup>Radboudumc, Department of Paediatric Neurology, Nijmegen, The Netherlands

<sup>4</sup>Paracelsus Medical University, Department of Paediatrics, Salzburg, Austria

<sup>5</sup>Université de Strasbourg, Centre National de la Recherche Scientifique, Illkirch, France

<sup>6</sup>Department of Clinical Genetics 4062, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark

<sup>7</sup>McMaster University, Department of Paediatrics, Hamilton, Canada

<sup>8</sup>Mitochondrial and neuromuscular diseases Laboratory '12 de Octubre' Hospital Research Institute. Centre for Biomedical Network Research on Rare Diseases, Madrid, Spain

<sup>9</sup>Mitochondrial Research Group, Newcastle University, Newcastle Upon Tyne, UK

<sup>10</sup>Women's and Children's Hospital, Adelaide, Australia

<sup>11</sup>Paediatric Neurology Unit, Hadassah University Hospital, Jerusalem, Israel

<sup>12</sup>Department of Genetics, University of Angers, France

<sup>13</sup>Laboratoire de Biochimie, APHP-CHU de Bicêtre, Cedex, France

Journal of Inherited Metabolic Disease 2012 Sep;35(5):737-47

Mitochondrial Complex I is the largest multi-protein enzyme complex of the oxidative phosphorylation system. Seven subunits of this complex are encoded by the mitochondrial and the remainder by the nuclear genome. We review the natural disease course and signs and symptoms of 130 patients (four new cases and 126 from literature) with mutations in nuclear genes encoding structural Complex I proteins or those involved in its assembly. Complex I deficiency caused by a nuclear gene defect is usually a non-dysmorphic syndrome, characterised by severe multi-system organ involvement and a poor prognosis. Age at presentation may vary, but is generally within the first year of life. The most prevalent symptoms include hypotonia, nystagmus, respiratory abnormalities, pyramidal signs, dystonia, psychomotor retardation or regression, failure to thrive, and feeding problems. Characteristic symptoms include brainstem involvement, optic atrophy, and Leigh syndrome on MRI, either or not in combination with internal organ involvement and lactic acidemia. Virtually all children ultimately develop Leigh syndrome or leukoencephalopathy. Twenty-five percent of the patients died before the age of six months, more than half before the age of two and 75% before the age of ten years. Some patients showed recovery of certain skills or are still alive in their thirties. No clinical, biochemical, or genetic parameters indicating longer survival were found. No clear genotype-phenotype correlations were observed, however defects in some genes seem to be associated with a better or poorer prognosis, cardiomyopathy, Leigh syndrome or brainstem lesions. longer rely on potentially biased data, with the associated costs and risks.

## Introduction

Mammalian Complex I or NADH:ubiquinone oxidoreductase is the largest enzyme complex of the mitochondrial oxidative phosphorylation system.<sup>1</sup> It consists of 45 subunits, of which seven are encoded by the mitochondrial genome, and the remainder by the nuclear genome.<sup>1</sup> Complex I has two modes of action: funnelling electrons to ubiquinone (co-enzyme Q) and redox driven proton translocation.<sup>2,3</sup> These actions are carried out by three proposed functional modules, consisting of several subunits of Complex I: the N-module for NADH oxidation, the Q-module for ubiquinone reduction and the P-module for proton translocation.<sup>4</sup> The complex has two arms, one embedded in the mitochondrial inner membrane (P-module) and one protruding into the mitochondrial matrix (N/Q-module), forming an L-shape. The proton gradient built up by Complex I to V is used by Complex V to synthesise ATP from ADP and inorganic phosphate. Fourteen highly conserved subunits can be distinguished, including the mitochondrial encoded *ND1-6*, *4L*, and *NDUFS1-3*, *NDUFS7-8* and *NDUFV1-2*.<sup>2</sup> The assembly of Complex I is not fully elucidated yet, but more and more assembly factors are found.<sup>6-8</sup>

Mutations in one of the nuclear encoded structural or assembly genes of Complex I have a dramatic effect on neurodevelopment and overall patient survival.<sup>9</sup> The majority of the children with an isolated Complex I deficiency present with Leigh syndrome, a devastating neurodegenerative disease.<sup>9-11</sup> These patients usually present within the first months of life with psychomotor retardation in combination with signs of brainstem or extrapyramidal dysfunction and lactic acidemia.<sup>10,12</sup> The disease has first been described by Denis Leigh who described the characteristic findings on neuropathological postmortem examination with vacuolation of the neuropil and relative preservation of the neurons, associated with demyelination, gliosis and capillary proliferation,<sup>13</sup> seen on MRI as bilateral hyperintensities in the basal ganglia, brainstem, thalamus, diencephalon, cerebellum and spinal cord.<sup>10</sup> Death occurs usually within the first years of life as a consequence of respiratory failure caused by ongoing brainstem dysfunction whether or not in combination with increasing muscle weakness.<sup>12</sup> Other phenotypes that have been described in patients with Complex I deficiency include neonatal cardiomyopathy,<sup>9,14-19</sup> leukoencephalopathy,<sup>16,20-22</sup> fatal infantile lactic acidosis<sup>11,23</sup> and other undefined progressive or stable encephalomyopathies.<sup>11,24</sup> The significance of the classification of symptoms into syndromes is still under debate.<sup>9</sup> No obvious genotype-phenotype correlations have been identified to date and patients with mutations in the same gene may present with highly variable phenotypes.<sup>9,25</sup> Also, the prognosis of nuclear encoded Complex I deficiency is quite variable, ranging from fatal neonatal disease<sup>11,23</sup> to survival beyond three decades.<sup>26</sup> For the patients and their families, it is important to get an evidence based indication of the prognosis for their child.

An accurate prognosis not only includes the age of death, but also the occurrence of symptoms such as epilepsy, visual disturbances, hearing problems, and brainstem symptoms, which may severely affect quality of life. The known presence of cardiomyopathy or deterioration during infection may have implications for follow-up and preventive immunisations. Importantly, the prediction of the clinical course of Complex I deficiency is not only important for the patients and their families, it is also indispensable for establishing clinical trials. Before the effect of a drug can be tested, the natural history must be known, with the most debilitating and most prevalent symptoms identified. To date, we are not aware of any treatments or supplementation with vitamins, anti-oxidant or other compounds that positively influence the disease course of these patients.<sup>27</sup>

In this review, we provide detailed information regarding the prognosis of patients with nuclear encoded Complex I deficiency based on systematic literature search. We summarise the clinical details of all nuclear encoded Complex I patients described in the literature, as well as four new cases with known mutations in Complex I genes. To give a detailed overview of prevalence of clinical symptoms and their natural course in time we looked for genetic, biochemical and clinical predictors indicative for prognosis of patients and based on this provide advice to clinicians taking care of Complex I deficient patients.

## Methods

### Search strategy for literature study

We searched Pubmed for all the individual subunits and assembly factors of Complex I to identify cases of nuclear encoded Complex I deficiency. We excluded patients with combined gene deletions. We contacted all research groups describing living patients with nuclear gene mutations and Complex I deficiency presented in the literature to assess the current clinical condition and the disease course after the publication. In patients on whom limited clinical information was present, we also contacted the authors to ask for a more detailed clinical description. Siblings of patients presented in the articles on whom no biochemical or genetic analysis was performed, were excluded from this review.

### New cases

We briefly described the clinical course of four new patients with known mutations in structural or assembly genes of Complex I.

### Genetic, biochemical and clinical data

We entered all data of the literature search and the new cases in a database, including age at presentation, age of death, cause of death, current age if the patient is still

alive, their clinical condition, the causative gene mutation, the origin of the patient, the presence of consanguinity, and histological findings in the muscle biopsy.

The activity of Complex I in skeletal muscle and skin fibroblasts was expressed as the percentage of the lowest reference value, to make the results of the measurements uniform, despite the different methods and reference values used in different laboratory. For the other patients, Complex I deficiency was confirmed by other methods such as Blue Native Page. The maximum lactate, both in serum and cerebrospinal fluid (CSF) was noted and considered (arbitrary) mildly elevated above 2.1 mmol/l (noted as 1 in the Supplementary Table), moderately elevated above 4.0 mmol/l (2) and severely elevated above 7.0 mmol/l (3). Alanine in serum was considered increased above 450  $\mu$ mol/l. The presence of tricarboxylic acid (TCA) cycle intermediates in urine was also noted.

Birth parameters were assessed, including the gestational age, birth weight and APGAR score. Low birth weight was defined as a birth weight lower than the 5<sup>th</sup> percentile for gestational age. The presence of dysmorphic features, macrocephaly, and microcephaly was noted. The age at which the following symptoms were described, was noted: psychomotor retardation, (isolated) motor retardation, developmental regression, deterioration after infection, failure to thrive, feeding problems, encephalopathy, lethargy, irritability, pyramidal signs and symptoms, extrapyramidal signs and symptoms, dys-/hypertonia, hypotonia, muscle weakness, exercise intolerance, dystrophy, ataxia, neuropathy, epilepsy, myoclonic epilepsy, ptosis, ophthalmology, nystagmus, strabismus, optic atrophy/pale optic disc, retinitis pigmentosa, vision problems, hearing loss, respiratory abnormalities, temperature regulation abnormalities, tension regulation abnormalities, dysphagia, cardiomyopathy, gastroesophageal reflux, vomiting, constipation, hepatopathy, renal involvement, and osteoporosis. The presence and localisation (basal ganglia, midbrain, brainstem, spinal cord, or cerebellum) of Leigh syndrome on brain magnetic resonance imaging (MRI) was noted, as well as the presence of cerebral/cerebellar atrophy, leukoencephalopathy, and hypoplasia of the corpus callosum. We noted the percentage of patients (literature and new cases) known to have the clinical features described above, as well as the median age at which the symptoms were first noted.

## Complex I subunits

Several sequential and parallel steps within the assembly process can be distinguished; the step in which the subunit is incorporated in the holocomplex was grouped according to Vogel *et al.*<sup>7</sup> The functional modules of Complex I as described in the introduction were grouped according to Angerer *et al.*<sup>4</sup> Brandt analysed whether the proteins were in the central core of the Complex or accessory to it; we grouped the subunits according to his analysis.<sup>2</sup>

### Statistical analysis:

All data were analysed using SPSS 16.0. Spearman's rho was used to correlate non-parametric data. All data were described using the median and the range. The number of patients dying before a certain age was calculated as number of patients died before that age divided by the total number of patients who died + patients living beyond that age. Outcome variables were tested for normality if the data reflected a Gaussian distribution test. Group medians were compared using the Mann-Witney test, or the Kruskal Wallis test. Correlations were calculated using the Spearman rank coefficient. Kaplan Meier survival analyses were performed for the patients with mutations in the three functional modules,<sup>4</sup> in assembly factors compared to structural genes, as well as for patients with mutations in central subunits or accessory subunits,<sup>2</sup> and in early and late in assembly.<sup>7</sup>

Since the number of patients per gene was insufficient to perform statistical pattern analyses, we evaluated the clinical course for specific pattern in the individual genes and individual symptoms if four patients or more with mutations in these genes are known.

### Case reports

The first patient was a boy with a homozygous deletion of exons 2-4 in *NDUFAF2* c.[128-?\_510+?del];[128-?\_510+?del]. He was the fourth child of first cousin parents of Lebanese descent. Three older siblings were healthy. Apart from maternal diabetes mellitus, the pregnancy was uncomplicated and he was born at 36 weeks gestation by Caesarean section with normal APGAR scores. He had bilateral single palmar creases. At six months of age, he was admitted to hospital due to a respiratory syncytial virus infection and psychomotor retardation was noted. Nystagmus was found, which had reportedly been present since age 4 months. An ophthalmological examination showed horizontal nystagmus, hypermetropia and decreased vision. In addition he had hearing impairment and used a hearing aid, and he was found to be developmentally delayed. At the age of nine months, he was admitted because of pneumonia. He developed epilepsy with generalised seizures and was treated with Phenobarbital; EEG was normal. Respiratory support was required due to apnoeas. He had episodes of sweating and pooling of secretions was also noted. Cerebral MRI showed bilateral symmetric signal changes in putamen, substantia nigra, cerebellar peduncles and brain stem. MR spectroscopy showed elevated lactate concentration of 10 mM. Soon after admission he died. In muscle, related to CS, the activity was 0.04 (reference range 0.20 - 0.54), and related to complex II: 0.15 (reference range 0.43 – 1.33). Complex I activity was not measured in fibroblasts.

The second patient, with a homozygous p.Val122Met *NDUFS7* mutation, was a son of healthy, non-consanguineous Dutch parents. He had normal development until 11 months, when he regressed and his growth stagnated. The patient started vomiting and severe gastro-oesophageal reflux was found. On physical examination, nystagmus, ophthalmoplegia and hypotonia were noted. Cranial MRI showed bilateral hyperintensities in the medulla oblongata, medial thalamus and cerebral penduncles. He died at the age of 20 months, after a viral infection with rapid neurological deterioration preceding hypoventilation and coma. Complex I activity in this patient was 353 mU/U CII (reference range 783 – 1,497 mU/U CII).

The third case is a boy with a homozygous p.Val122Met *NDUFS7* mutation, born from healthy, non-consanguineous Dutch parents. He developed normally until 9 months, when parents noticed he was clumsy but able to learn to walk. At one year of age, nystagmoid eye movements were observed. At the age of 2 years and 3 months, he was admitted to a local hospital with progressive gait disturbances, lethargy and articulation problems. No laboratory abnormalities were observed. At physical examination two months later, he had a bilateral ptosis, hypertonia of all limbs, severe axial ataxia, intention tremor and hypertension. A slightly elevated lactate was found in serum and CSF fluid. MRI showed hyperintensities in the medial thalamus, brainstem and cerebellum, indicative of Leigh syndrome. One month later, he developed respiratory insufficiency, for which he was admitted to the intensive care, but continued to deteriorate. He is now 10 years old and severely disabled due to contractures and dystonia, but able to make good contact with his environment. An NADH : ubiquinone oxidoreductase activity of 24 mU/U CS (reference range 70 – 250 mU/U CS) was found in skeletal muscle.

The fourth patient, with a heterozygous *NDUFV1* p.Ser56Pro and p.Thr423Met mutation, was born from healthy, non-consanguineous parents. His development was normal until the age of 8 months, when he deteriorated after an ear infection. His development regressed until the age of 11 months, followed by a partial recovery of motor skills. MRI showed white signal abnormalities, without involvement of the basal ganglia and brainstem. At present, he shows motor delay, including pyramidal signs, hypotonia and a mild ataxia, but is able to walk with support. He is 2.5 years old with minor cognitive impairment with borderline-normal language production. The NADH:ubiquin oxidoreductase activity was 50 mU/U CS in skeletal muscle (reference range 100 – 401 mU/U CS).

## Results

### Search results

Forty-eight articles describing patients with nuclear encoded Complex I deficiency were found by searching the Pubmed database,<sup>8,17,20-22,26,28-68</sup> identifying a total of 126 patients. We additionally describe four new Complex I deficient cases of nuclearDNA origin. Ten out of twelve colleagues responded to our request for more information on patients who were still alive when the article was written or on patients who had a incomplete clinical case description.

### Genetic background

Including the new cases, the assembly factor group consists of 44 patients: two patients with a mutation in *NDUFAF1*, eight patients with a mutation in *NDUFAF2*, four patients with a mutation in *NDUFAF3*, nine patients with mutations in *NDUFAF4*, seven patients with mutations in *ACAD9*, two patients with mutations in *FOXRED1*, one patient with a mutation in *NUBPL*, eight patients with a mutation in *C20orf7*, and two patients with mutations in *C8orf38*.

The structural gene group consists of: five patients with a mutation in *NDUFA1*, one patient with a mutation in *NDUFA2*, one patient with a mutation in *NDUFA10*, five patients with a mutation in *NDUFA11*, one patient with a mutation in *NDUFA12*, ten patients with mutations in *NDUFS1*, fifteen patients with a mutation in *NDUFS2*, one patient with a mutation in *NDUFS3*, fourteen patients with a mutation in *NDUFS4*, six patients with a mutation in *NDUFS6*, six patients with a mutation in *NDUFS7*, three patients with a mutation in *NDUFS8*, seventeen patients with a mutation in *NDUFV1* and one patient with a mutation in *NDUFV2*. No patients with mutations in *NDUFV3*, *NDUFS5*, *NDUFA3-9*, *NDUFA12-13*, *NDUFC1-2*, *NDUFB1-11* or *NDUFAB1* were found in literature. We additionally described four patients with mutations in *NDUFAF2*, *NDUFS7*, and *NDUFV1*. Sixty children (71%) were born from consanguineous parents. Fifty-six patients belonged to 23 families, the other patients were the only patient within the family.

### Biochemical results

Median complex I activity in muscle was 29% of the lowest normal reference value, with a range of 3 - 100% ( $n = 68$ ). The three patients with normal Complex I activity in muscle had a low Complex I activity in fibroblasts. Complex I activity in fibroblasts had a median value of 35% of the lowest normal reference value, with a range of 5 - 82% ( $n = 59$ ). Of the 14 patients reported to have abnormal findings on histology, three patients had ragged red fibres, seven had increased lipid content of the muscle, two had atrophic type 2 fibres, one had a atrophic of type 1 fibres, four had abnormal mitochondria, and two had a reduced number of type 1 fibres. In five patients, excretion



of TCA cycle intermediates was reported ( $n = 12$ ). Of the 96 patients in which lactate levels were reported, 86 had an increased lactate, of which 37 had a mildly increased lactate, 20 had a moderately increased lactate, and 29 had a severely increased lactate concentration. Lactate in CSF was elevated in 31 out of 34 patients of which 14 had a mild increase and 15 had a moderate increase. Serum alanine was only reported in 9 patients, all but one being increased.

### Clinical details

See Figure 1, Table 1 and the Supplementary Document 5 for a detailed description of the age of presentation and age of death of patients, categorised by gene mutation, including the new cases described. Sixty-seven boys (58%) and 49 girls were reported; of 15 patients no gender was mentioned in the case report. The median age at presentation was four months (range 0 months to nine years;  $n = 130$ ). Median age of death was 10 months (range 0 months to 13.5 years;  $n = 90$ ). Thirty-four patients died before the age of six months (25%), 47 died before the age of one year (36%), 74 patients died before the age of two years (58%), 81 patients died before the age of four years (66%) and 85 patients died before the age of ten years (75%). The patients who are still alive are now six months to 38 years old (median age 9 years;  $n = 33$ ). Nine patients were reported to have died from a cardiorespiratory failure, fifteen patients died from respiratory failure, six patients died from central hypoventilation, three patients died from multiple organ failure, twenty-one patients died from lactic acidosis, two patients suffered from aspiration pneumonia, three patients from cardiomyopathy, and four patients died from infection.



**Table 1.****Age of death.** Age of death (range, in years) per mutation; > means beyond the age of.

Gene mutated	Range age of death (years)	Number of observations
<i>NDUFA1</i>	0.6 - >24	2
<i>NDUFA2</i>	0.8 - 13.5	9
<i>NDUFA3</i>	0.2 - 0.3	4
<i>NDUFA4</i>	0 - 1.5	9
<i>ACAD9</i>	0.1 - >12	7
<i>FOXRED1</i>	>10 - >22	2
<i>NUBPL</i>	>8	1
<i>C20orf7</i>	0 - >29	8
<i>C8orf38</i>	> 1.8 - 2.9	2
<i>NDUFA1</i>	1.2 - >38	5
<i>NDUFA2</i>	0.9	1
<i>NDUFA10</i>	1.9	1
<i>NDUFA11</i>	0 - 4	5
<i>NDUFA12</i>	> 10	1
<i>NDUFS1</i>	0.4 - >4	10
<i>NDUFS2</i>	0.3 - >17	15
<i>NDUFS3</i>	13	1
<i>NDUFS4</i>	0.3 - 2.3	14
<i>NDUFS6</i>	0	6
<i>NDUFS7</i>	0.5 - 5	6
<i>NDUFS8</i>	0.2 - >9	3
<i>NDUFV1</i>	0.3 - >13	17
<i>NDUFV2</i>	0.2	1

In seven patients, a microcephaly was described, one patient had macrocephaly with wide anterior fontanel, in one patient a hydroureter with hydronephrosis and in one child upslanting palpebral fissures and hypospadias were reported. The latter patient was a child of consanguineous parents. For a more detailed summary of the prevalence and age of presentation of the individual symptoms (Table 2).

**Table 2.**

**Reported clinical symptoms in patients with nuclear encoded Complex I deficiency.** Reported symptoms including the prevalence and mean age of presentation in our cohort ( $n = 130$ ).

Clinical symptom	Prevalence (%)	Median age of onset (months)	Range (years)
Hypotonia	60	5	0 - 6
Failure to thrive	34	6	0 - 2.8
Nystagmus	34	7	0 - 10
Dys- or hypertonia	32	12	0 - 9
Psychomotor retardation	30	6	0 - 6
Feeding problems	29	5	0 - 5.2
Pyramidal symptoms	28	13	0 - 13
Respiratory abnormalities	27	11	0 - 12
Developmental regression	25	11	0 - 7.5
Vomiting	22	7	0 - 4
Epilepsy	21	8	0 - 10
Cardiomyopathy	20	4	0 - 3
Optic atrophy	20	10	0 - 11
Deterioration after infections	18	6	0.3 - 3.2
Ataxia	18	24	0.5 - 8
Lethargy	18	6	0 - 20
Encephalopathy	16	6	0 - 2.5
Vision problems	16	6	0 - 10
Muscle weakness	15	8	0 - 11.6
Irritability	15	7	0 - 16
Extrapyramidal symptoms	15	14	0.3 - 10
Strabismus	14	7	0.1 - 6
Dysphagia	13	6	0 - 9
Pure motor retardation	11	9	0 - 2
Dystrophy	9	6	0.3 - 8
Myoclonic epilepsy	8	10	0.2 - 10
Hearing loss	8	16	0.1 - 10
Ptosis	7	16	0.3 - 2.5
Exercise intolerance	7	48	0 - 20
Temperature regulation problems	5	6	0.4 - 20
Gastrooesophageal reflux	5	6	0 - 1.1
Hepatopathy	4	4	0 - 1.3
Ophthalmoplegia	4	9	0.6 - 2.2
Constipation	4	6	0.2 - 7
Osteoporosis	3	78	5-16

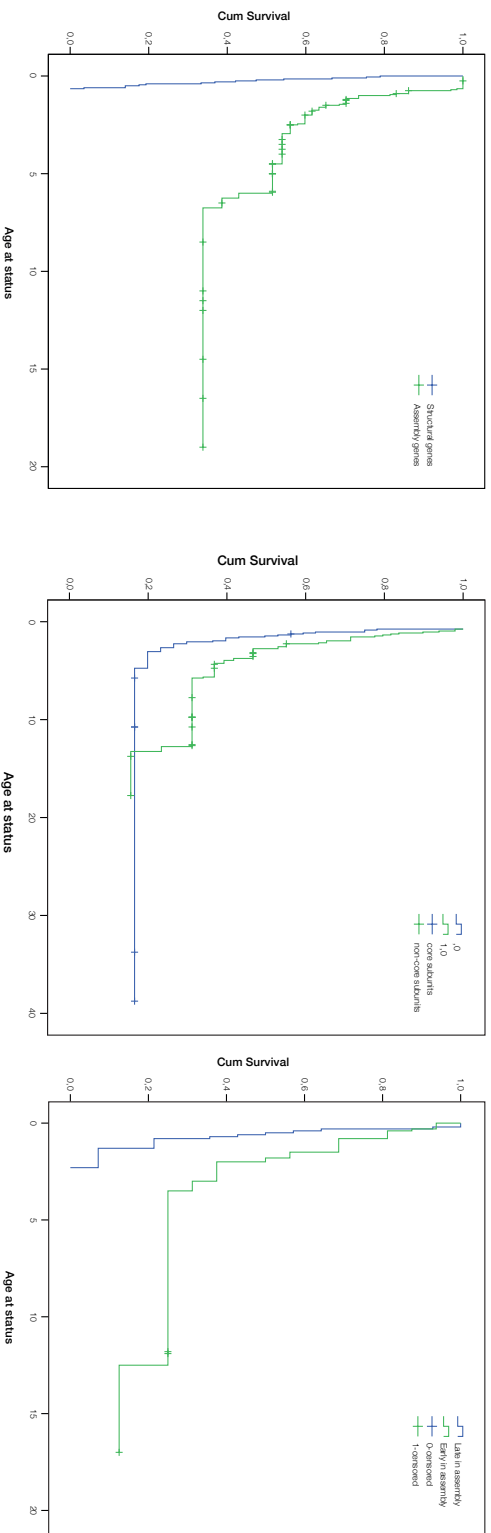
Neuropathy	3	8	0.5 - 30
Tension abnormalities	2	51	0.5 - 8
Renal involvement	2	17	0.5 - 2
Retinitis pigmentosa	1		13

Sixty-five patients with Leigh syndrome were described, of which 47 had hyperintensities in the basal ganglia and thalamus, 14 had lesions in the midbrain, and 23 had lesions in the brainstem. Spinal cord hyperintensities were described in one patient and lesions in the cerebellum were reported in four patients. Fourteen patients had cerebral atrophy and six patients had cerebellar atrophy. Leukoencephalopathy was present in 26 patients. Hypoplasia of the corpus callosum was present in three patients. In two patients, no abnormalities on brain MRI were described. For a detailed description of the MRI abnormalities in our patients (Table 3). Of the 50 patients who had lesions in the basal ganglia or midbrain, 27 had pyramidal or extrapyramidal features (54%). Of the 23 patients with hyperintensities in the brainstem, 20 had features of brainstem dysfunction (87%) and 13 had pyramidal or extrapyramidal signs (56%). Of the 24 patients with leukoencephalopathy, 17 had pyramidal or extrapyramidal signs (70%).

**Table 3.**

**The prevalence of MRI abnormalities in patients with nuclear encoded Complex I deficiency.**

MRI abnormality	Prevalence (%)
Leigh syndrome	84
basal ganglia	53
midbrain	17
brainstem	27
spinal cord	1
cerebellum	4
Cerebral atrophy	12
Cerebellar atrophy	9
Leukoencephalopathy	28
Hypoplasia of the corpus callosum	3
Normal MRI	2



**Figure 2.**

**Survival curve of Complex patients \*** Age in years (x-axis) and cumulative survival (y-axis). A) Survival of patients with mutations in assembly genes (green) compared to patients with mutations in structural genes (blue) ( $b = 5.2$ ;  $HR = 186$ ;  $p < 0.001$ ;  $n = 123$ ). B) Survival of patients with mutations in core subunits (blue) compared to patients with mutations in non-core genes (green) ( $b = -0.6$ ;  $HR = 0.55$ ;  $p = 0.022$ ;  $n = 81$ ). C) Survival of patients with mutations in genes encoding proteins built late in the assembly was worse compared to survival of patients with mutations in genes encoding subunits which were built in early in the assembly ( $b = 0.95$ ;  $HR 0.39$ ;  $p = 0.029$ ;  $n = 30$ )

\*Erratum to original figure

## Statistical analysis

No correlation between the Complex I activity and age of onset ( $p = 0.38$ ;  $n = 82$ ) or age of death ( $p = 0.15$ ;  $n = 53$ ) was found. A moderate correlation was found between age of presentation and age of death ( $r = 0.58$ ;  $p < 0.001$ ;  $n = 91$ ). No significant differences in age of death or age of onset were found between the functional subunits, order within the assembly of the holo-complex, or the subcomplex in which the protein is present. Survival of patients with mutations in assembly genes was lower compared to patients with mutations in structural genes ( $b = 5.2$ ; HR = 186;  $p < 0.001$ ;  $n = 123$ ). Survival of patients with mutations in core subunits was lower than of patients with mutations in non-core genes ( $b = -0.6$ ; HR = 0.55;  $p = 0.022$ ;  $n = 81$ ). Survival of patients with mutations in genes encoding proteins built late in the assembly was worse compared to survival of patients with mutations in genes encoding subunits which were built in early in the assembly ( $b = 0.95$ ; HR 0.39;  $p = 0.029$ ;  $n = 30$ ). No difference could be observed between patients having mutations in different functional subunits ( $p = 0.775$ ; Figure 2).

## Discussion

In this review, we describe the disease course of 130 patients with mutations in either structural or assembly proteins leading to Complex I deficiency. The disease course of nuclear encoded Complex I deficiency is quite homogeneous, with ultimately most children having a severe multi-system disease with prominent neurological involvement. Most children with an age of younger than six months presented with hypotonia, feeding problems and failure to thrive, vomiting, encephalopathy, epilepsy, and eye movement disturbances. Children beyond the age of six months more frequently presented with psychomotor retardation or developmental regression, pyramidal signs and symptoms, dystonia, ataxia, epilepsy, failure to thrive, vomiting, and optic atrophy. The most specific signs pointing to a nuclear encoded Complex I deficiency are brainstem involvement, optic atrophy and Leigh syndrome characteristics on MRI. Cardiac, renal, or hepatic involvement was seen in 24% of the children. Deterioration with infections was described in only 18% of the patients, which is less frequent than generally found in children with mitochondrial disorders and may therefore be an underreporting. No dysmorphic features were present, except for two children of consanguineous parents. Most children appear normal at birth, five percent of the patients were born mildly premature and seven percent had a low birth weight, similar to what was previously described.<sup>9</sup> Increased lactic acid concentration was present in 90% of the patients.

Importantly, not all patients with a nuclear encoded Complex I deficiency have a poor prognosis and quality of life. For example the girl with a compound heterozygous *NDUFV1* mutation described by Zafeiriou *et al.*<sup>21</sup> has developed normally after the initial regression around the age of one year. She has learning problems and mild spasticity,

but is able to ride a bicycle and walk without support. The patient with a compound heterozygous mutation in *NDUFAF1*, described by Dunning *et al.*,<sup>30</sup> is still alive at the age of 24 years. He works part time, he has Asperger syndrome and kyphosis, but has a sense of humour is generally cheerful.

Only two patients without abnormalities on MRI were reported.<sup>31</sup> Possibly, these siblings would have developed brain abnormalities if they had lived longer than three months. Of the patients dying before the age of six months, two patients were reported without any brain abnormalities (15%), whereas only four patients older than six months (6%) were described with brain atrophy without Leigh syndrome or leukoencephalopathy. No clear correlation between the anatomic location of the hyperintensities and the clinical symptoms was observed.

Historically, no obvious mutation-related phenotype apparent in Complex I deficiency was reported.<sup>9</sup> For example, cardiomyopathy has first been reported as characteristic for the *NDUFS2* gene<sup>39</sup> but was later reported in other nuclear Complex I genes as well.<sup>9</sup> Moreover, after the first three patients with *NDUFS2* mutations with cardiomyopathy, no patients with this combination have been described (Supplementary Document 5).<sup>65</sup> Besides, mutations in the same gene can express a high variety of clinical and biochemical phenotypes, even within the same family.<sup>42</sup> If more than four patients with mutations in a certain gene have been described we analysed the genotype-phenotype relation. We found that cardiomyopathy was more prevalent in patients with *ACAD9* and *NDFUA11* mutations. In patients with *NDUFAF2* mutations, brainstem symptoms and Leigh syndrome on MRI were observed in all patients, as well as a high prevalence of pale optic discs or optic atrophy. Patients with *NDUFAF4* and *NDUFS6* patients seem to have a very poor prognosis, but since all *NDUFAF4* patients are from the same family, no solid conclusions can be drawn. In patients with *NDUFA1* mutations, hearing loss seems more common and prognosis is often quite good. Leukoencephalopathy seems more prevalent in patients with *NDUFS1* and *NDUFV1* mutations, but mutations in both genes may also lead to Leigh syndrome. Patients with *NDUFS4* and *NDUFS7* mutations almost invariably have Leigh syndrome, in patients with *NDUFS7* mutations, the brainstem is often affected.

The prognosis of nuclear Complex I deficiency is generally very poor (Figure 1 and 2), however quite a few exceptions exist. More than half of the patients died before the age of two years and 79% died before the age of ten years. In comparison, children with a mitochondrial DNA encoded Complex I deficiency present at the median age of 12 months and have a tendency towards a longer survival than patients with nuclear encoded Complex I deficiency.<sup>69</sup> Although lactic acidosis was described as the cause of death in 41% of the patients dying before the age of six months, a high lactate is not necessarily associated with a poor prognosis. The same applies to a high lactate in



the CSF. No correlation between the age of death and the location of the protein within the assembly, function or structure of Complex I could be elucidated. Strikingly, within one family harbouring the same mutation, one child may die before the age of one year whereas a sibling lives into the first decade.<sup>17</sup> In the patients living into their first decade, no lower prevalence of cardiomyopathy, brainstem dysfunction, Leigh syndrome, or a higher Complex I activity could be observed. Although patients with mutations in core subunits have a significantly poorer prognosis than patients with mutations in non-core subunits, the variability in prognosis is extremely high. The value of predicting outcome based on this kind of molecular characteristics warrants further research, *e.g. in silico*.

In previous studies in a heterogeneous group of children with mitochondrial disorders, cardiomyopathy was found to be a significant predictor of a fatal outcome.<sup>70</sup> However, since the neurological phenotype of patients with nuclear encoded Complex I deficiency is so overwhelming, this observation is difficult to confirm in our cohort. For example, all patients with mutations in Complex I assembly genes and cardiomyopathy survived beyond the age of five years, although all children with mutations in structural genes and cardiomyopathy died before the age of two years. Cardiomyopathy may be present in both patients with structural and assembly factor mutations, and in combination with various clinical phenotypes ranging from lactic acidemia and progressive encephalomyopathy,<sup>39</sup> to mild and stable neurological syndromes.<sup>30,62</sup>

Although cardiomyopathy always presented before the age of three years in this cohort, cardiac function should be checked regularly in all patients with nuclear encoded Complex I deficiency. We also advice to regularly check the ocular manifestations of the disease, for it is important to adapt the environment to visual disturbances in children with optic neuropathy, severe ptosis or ophthalmoplegia.<sup>71</sup> General advices, such as tube-feeding in case of malnutrition or aspiration, and aggressively treating fever and infectious diseases also apply to these patients to prevent secondary malfunction or challenging of the energy metabolism.<sup>72</sup> For patients with dystonia or spasticity, a combination of physiotherapy and pharmacotherapy may relieve symptoms.<sup>71</sup>

This is the first retrospective study of the clinical disease course of patients with a nuclear encoded Complex I deficiency. Although the number of patients analysed is high considering the prevalence of this condition, the number of patients with particular gene mutations is still too low to see clear genotype-phenotype patterns. Besides, an obvious report bias exists and the quality of our data is inherently limited by the quality and completeness of the data reported by other research groups. Some case reports were very brief and only reported the main symptoms. Furthermore, it is more likely that the symptoms present in previously reported patients with the same mutation will be diagnosed and reported, *e.g.* optic atrophy or retinitis pigmentosa. The median and mean age we calculated was based on the observation and description of others,

sometimes in only a few papers. Therefore, the numbers in our paper should be carefully interpreted and interpreted in a clinical context.

For future research, we would suggest to perform prospective follow-up of these patients, for example with the Newcastle Paediatric Mitochondrial Disease Scale.<sup>73</sup> Only a prospective follow-up will provide valid data to be used in the preparation clinical trials in predicting the natural disease course or selecting relevant outcome measures.

We also suggest to continue publishing clinical details of these rare diseases. Only more case descriptions will enable us to predict the natural disease course of patients with these rare diseases, which is not only useful for future clinical trials, but is also indispensable for the patients and their families.

## Conclusion

In conclusion, nuclear encoded Complex I deficiency generally has a devastating clinical disease course, characterised by a severe neurological phenotype including brainstem involvement and optic atrophy, in combination with lactic acidosis and Leigh syndrome on MRI. Most patients die before the age of one year, but patients in their third decade have been described. We advise to regularly check cardiac and ocular manifestations of the disease, to optimise nutrition, to treat intercurrent illnesses promptly, and to provide adequate symptomatic relieve and family support.

## References

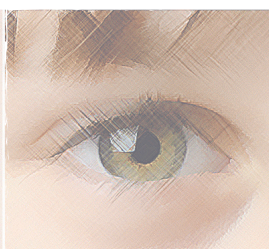
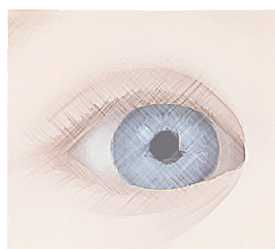
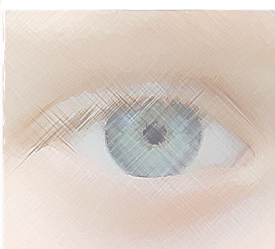
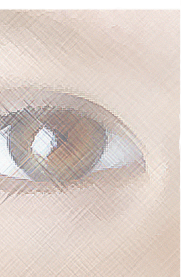
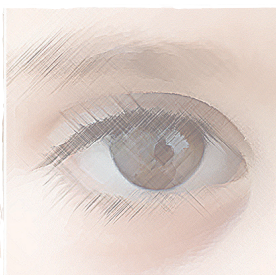
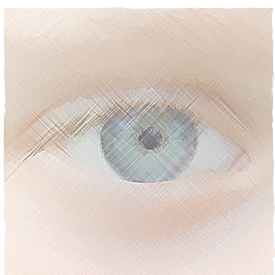
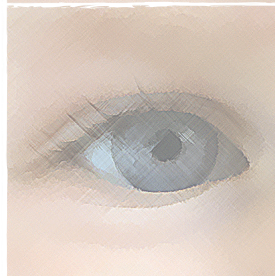
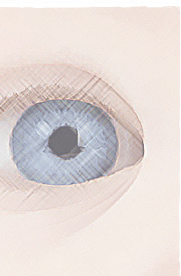
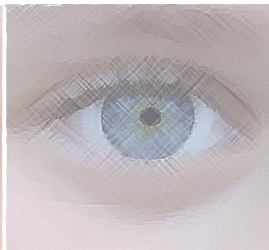
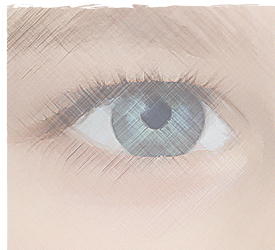
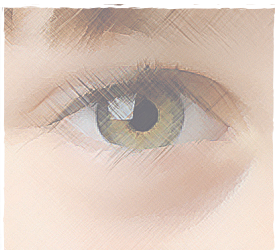
1. Carroll J, Fearnley IM, Skehel JM, Shannon RJ, Hirst J, Walker JE. Bovine complex I is a complex of 45 different subunits. *The Journal of biological chemistry* 2006;281:32724-7.
2. Brandt U. Energy converting NADH:quinone oxidoreductase (complex I). *Annual review of biochemistry* 2006;75:69-92.
3. Efremov RG, Baradaran R, Sazanov LA. The architecture of respiratory complex I. *Nature* 2010;465:441-5.
4. Angerer H, Zwicker K, Wumaier Z, *et al.* A scaffold of accessory subunits links the peripheral arm and the distal proton-pumping module of mitochondrial complex I. *The Biochemical journal* 2011;437:279-88.
5. Clason T, Ruiz T, Schagger H, *et al.* The structure of eukaryotic and prokaryotic complex I. *J Struct Biol* 2010;169:81-8.
6. McKenzie M, Ryan MT. Assembly factors of human mitochondrial complex I and their defects in disease. *IUBMB Life* 2010;62:497-502.
7. Vogel RO, Smeitink JA, Nijtmans LG. Human mitochondrial complex I assembly: a dynamic and versatile process. *Biochimica et biophysica acta* 2007;1767:1215-27.
8. Nouws J, Nijtmans L, Houten S, *et al.* Acyl-CoA dehydrogenase 9 is required for the biogenesis of oxidative phosphorylation complex I. *Cell metabolism* 2010;In press.
9. Distelmaier F, Koopman WJ, van den Heuvel LP, *et al.* Mitochondrial complex I deficiency: from organelle dysfunction to clinical disease. *Brain* 2009;132:833-42.
10. Rahman S, Blok RB, Dahl HH, *et al.* Leigh syndrome: clinical features and biochemical and DNA abnormalities. *Annals of neurology* 1996;39:343-51.
11. Loeffen JL, Smeitink JA, Trijbels JM, *et al.* Isolated complex I deficiency in children: clinical, biochemical and genetic aspects. *Human mutation* 2000;15:123-34.
12. Finsterer J. Leigh and Leigh-like syndrome in children and adults. *Paediatr Neurol* 2008;39:223-35.
13. Leigh D. Subacute necrotizing encephalomyelopathy in an infant. *Journal of neurology, neurosurgery, and psychiatry* 1951;14:216-21.
14. Janssen RJ, Nijtmans LG, van den Heuvel LP, Smeitink JA. Mitochondrial complex I: structure, function and pathology. *Journal of inherited metabolic disease* 2006;29:499-515.
15. Janssen AJ, Trijbels FJ, Sengers RC, *et al.* Measurement of the energy-generating capacity of human muscle mitochondria: diagnostic procedure and application to human pathology. *Clinical chemistry* 2006;52:860-71.
16. Bugiani M, Invernizzi F, Alberio S, *et al.* Clinical and molecular findings in children with complex I deficiency. *Biochimica et biophysica acta* 2004;1659:136-47.
17. Saada A, Edvardson S, Rapoport M, *et al.* C6ORF66 is an assembly factor of mitochondrial complex I. *American journal of human genetics* 2008;82:32-8.
18. Hoefs SJ, Dieteren CE, Distelmaier F, *et al.* NDUFA2 complex I mutation leads to Leigh disease. *American journal of human genetics* 2008;82:1306-15.
19. Wispe JR, Warner BB, Clark JC, *et al.* Human Mn-superoxide dismutase in pulmonary epithelial cells of transgenic mice confers protection from oxygen injury. *The Journal of biological chemistry* 1992;267:23937-41.

20. Benit P, Chretien D, Kadhom N, *et al.* Large-scale deletion and point mutations of the nuclear NDUFV1 and NDUF51 genes in mitochondrial complex I deficiency. *American journal of human genetics* 2001;68:1344-52.
21. Zafeiriou DI, Rodenburg RJ, Scheffer H, *et al.* MR spectroscopy and serial magnetic resonance imaging in a patient with mitochondrial cystic leukoencephalopathy due to complex I deficiency and NDUFV1 mutations and mild clinical course. *Neuropaediatrics* 2008;39:172-5.
22. Hoefs SJ, Skjeldal OH, Rodenburg RJ, *et al.* Novel mutations in the NDUF51 gene cause low residual activities in human complex I deficiencies. *Molecular genetics and metabolism* 2010;100:251-6.
23. Bentlage HA, Wendel U, Schagger H, ter Laak HJ, Janssen AJ, Tjibbels JM. Lethal infantile mitochondrial disease with isolated complex I deficiency in fibroblasts but with combined complex I and IV deficiencies in muscle. *Neurology* 1996;47:243-8.
24. Pitkanen S, Feigenbaum A, Laframboise R, Robinson BH. NADH-coenzyme Q reductase (complex I) deficiency: heterogeneity in phenotype and biochemical findings. *Journal of inherited metabolic disease* 1996;19:675-86.
25. Tucker EJ, Compton AG, Calvo SE, Thorburn DR. The molecular basis of human complex I deficiency. *IUBMB Life* 2011.
26. Potluri P, Davila A, Ruiz-Pesini E, *et al.* A novel NDUF41 mutation leads to a progressive mitochondrial complex I-specific neurodegenerative disease. *Molecular genetics and metabolism* 2009;96:189-95.
27. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med* 2012;366:1132-41.
28. Mayr JA, Bodamer O, Haack TB, *et al.* Heterozygous mutation in the X chromosomal NDUF41 gene in a girl with complex I deficiency. *Molecular genetics and metabolism* 2011;103:358-61.
29. Pagniez-Mammeri H, Landrieu P, Legrand A, Slama A. Leukoencephalopathy with vanishing white matter caused by compound heterozygous mutations in mitochondrial complex I NDUF51 subunit. *Molecular genetics and metabolism* 2010;101:297-8.
30. Dunning CJ, McKenzie M, Sugiana C, *et al.* Human CIA30 is involved in the early assembly of mitochondrial complex I and mutations in its gene cause disease. *The EMBO journal* 2007;26:3227-37.
31. Saada A, Vogel RO, Hoefs SJ, *et al.* Mutations in NDUF43 (C3ORF60), encoding an NDUF44 (C6ORF66)-interacting complex I assembly protein, cause fatal neonatal mitochondrial disease. *American journal of human genetics* 2009;84:718-27.
32. Ogilvie I, Kennaway NG, Shoubbridge EA. A molecular chaperone for mitochondrial complex I assembly is mutated in a progressive encephalopathy. *J Clin Invest* 2005;115:2784-92.
33. Sugiana C, Pagliarini DJ, McKenzie M, *et al.* Mutation of C20orf7 disrupts complex I assembly and causes lethal neonatal mitochondrial disease. *American journal of human genetics* 2008;83:468-78.
34. Pagliarini DJ, Calvo SE, Chang B, *et al.* A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 2008;134:112-23.
35. Fernandez-Moreira D, Ugalde C, Smeets R, *et al.* X-linked NDUF41 gene mutations associated with mitochondrial encephalomyopathy. *Annals of neurology* 2007;61:73-83.
36. Berger I, Hershkovitz E, Shaag A, Edvardson S, Saada A, Elpeleg O. Mitochondrial complex I deficiency caused by a deleterious NDUF411 mutation. *Annals of neurology* 2008;63:405-8.

37. Barghuti F, Elian K, Gomori JM, *et al.* The unique neuroradiology of complex I deficiency due to NDUFA12L defect. *Molecular genetics and metabolism* 2008;94:78-82.
38. Martin MA, Blazquez A, Gutierrez-Solana LG, *et al.* Leigh syndrome associated with mitochondrial complex I deficiency due to a novel mutation in the NDUFS1 gene. *Arch Neurol* 2005;62:659-61.
39. Loeffen J, Elpeleg O, Smeitink J, *et al.* Mutations in the complex I NDUFS2 gene of patients with cardiomyopathy and encephalomyopathy. *Annals of neurology* 2001;49:195-201.
40. Benit P, Beugnot R, Chretien D, *et al.* Mutant NDUFV2 subunit of mitochondrial complex I causes early onset hypertrophic cardiomyopathy and encephalopathy. *Human mutation* 2003;21:582-6.
41. Budde SM, van den Heuvel LP, Janssen AJ, *et al.* Combined enzymatic complex I and III deficiency associated with mutations in the nuclear encoded NDUFS4 gene. *Biochemical and biophysical research communications* 2000;275:63-8.
42. Budde SM, van den Heuvel LP, Smeets RJ, *et al.* Clinical heterogeneity in patients with mutations in the NDUFS4 gene of mitochondrial complex I. *Journal of inherited metabolic disease* 2003;26:813-5.
43. Papa S, Scacco S, Sardanelli AM, *et al.* Mutation in the NDUFS4 gene of complex I abolishes cAMP-dependent activation of the complex in a child with fatal neurological syndrome. *FEBS letters* 2001;489:259-62.
44. Leshinsky-Silver E, Lebre AS, Minai L, *et al.* NDUFS4 mutations cause Leigh syndrome with predominant brainstem involvement. *Molecular genetics and metabolism* 2009;97:185-9.
45. Anderson SL, Chung WK, Frezzo J, *et al.* A novel mutation in NDUFS4 causes Leigh syndrome in an Ashkenazi Jewish family. *Journal of inherited metabolic disease* 2008;Supplement 2:S461-7.
46. Petruzzella V, Vergari R, Puzziferri I, *et al.* A nonsense mutation in the NDUFS4 gene encoding the 18 kDa (AQDQ) subunit of complex I abolishes assembly and activity of the complex in a patient with Leigh-like syndrome. *Human molecular genetics* 2001;10:529-35.
47. Spiegel R, Shaag A, Mandel H, *et al.* Mutated NDUFS6 is the cause of fatal neonatal lactic acidemia in Caucasus Jews. *Eur J Hum Genet* 2009;17:1200-3.
48. Kirby DM, Salemi R, Sugiana C, *et al.* NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency. *J Clin Invest* 2004;114:837-45.
49. Lebon S, Minai L, Chretien D, *et al.* A novel mutation of the NDUFS7 gene leads to activation of a cryptic exon and impaired assembly of mitochondrial complex I in a patient with Leigh syndrome. *Molecular genetics and metabolism* 2007;92:104-8.
50. Lebon S, Rodriguez D, Bridoux D, *et al.* A novel mutation in the human complex I NDUFS7 subunit associated with Leigh syndrome. *Molecular genetics and metabolism* 2007;90:379-82.
51. Triepels RH, van den Heuvel LP, Loeffen JL, *et al.* Leigh syndrome associated with a mutation in the NDUFS7 (PSST) nuclear encoded subunit of complex I. *Annals of neurology* 1999;45:787-90.
52. Loeffen J, Smeitink J, Triepels R, *et al.* The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. *American journal of human genetics* 1998;63:1598-608.
53. Procaccio V, Wallace DC. Late-onset Leigh syndrome in a patient with mitochondrial complex I NDUFS8 mutations. *Neurology* 2004;62:1899-901.
54. Schuelke M, Smeitink J, Mariman E, *et al.* Mutant NDUFV1 subunit of mitochondrial complex I causes leukodystrophy and myoclonic epilepsy. *Nature genetics* 1999;21:260-1.

55. Brenningstall GN, Shoffner J, Patterson RJ. Siblings with leukoencephalopathy. *Semin Paediatr Neurol* 2008;15:212-5.
56. Laugel V, This-Bernd V, Cormier-Daire V, Speeg-Schatz C, de Saint-Martin A, Fischbach M. Early-onset ophthalmoplegia in Leigh-like syndrome due to NDUFV1 mutations. *Paediatr Neurol* 2007;36:54-7.
57. Saada A, Edvardson S, Shaag A, *et al.* Combined OXPHOS complex I and IV defect, due to mutated complex I assembly factor C20ORF7. *J Inherit Metab Dis.* 2012 Jan;35(1):125-31.
58. Fassone E, Duncan AJ, Taanman JW, *et al.* FOXRED1, encoding an FAD-dependent oxidoreductase complex-I-specific molecular chaperone, is mutated in infantile-onset mitochondrial encephalopathy. *Human molecular genetics* 2010;19:4837-47.
59. Hoefs SJ, Dieteren CE, Rodenburg RJ, *et al.* Baculovirus complementation restores a novel NDUF2 mutation causing complex I deficiency. *Human mutation* 2009;30:E728-36.
60. Calvo SE, Tucker EJ, Compton AG, *et al.* High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nature genetics* 2010;42:851-8.
61. Gerards M, van den Bosch BJ, Danhauser K, *et al.* Riboflavin-responsive oxidative phosphorylation complex I deficiency caused by defective ACAD9: new function for an old gene. *Brain* 2011;134:210-9.
62. Haack TB, Danhauser K, Haberberger B, *et al.* Exome sequencing identifies ACAD9 mutations as a cause of complex I deficiency. *Nature genetics* 2010;42:1131-4.
63. van den Heuvel L, Ruitenbeek W, Smeets R, *et al.* Demonstration of a new pathogenic mutation in human complex I deficiency: a 5-bp duplication in the nuclear gene encoding the 18-kD (AQDQ) subunit. *American journal of human genetics* 1998;62:262-8.
64. Hoefs SJ, van Spronsen FJ, Lenssen EW, *et al.* NDUF10 mutations cause complex I deficiency in a patient with Leigh disease. *Eur J Hum Genet* 2011;19:270-4.
65. Tuppen HA, Hogan VE, He L, *et al.* The p.M292T NDUF2 mutation causes complex I-deficient Leigh syndrome in multiple families. *Brain* 2010;133:2952-63.
66. Vilain C, Rens C, Aebi A, *et al.* A novel NDUFV1 gene mutation in complex I deficiency in consanguineous siblings with brainstem lesions and Leigh syndrome. *Clin Genet* 2012 Sep;82(3):264-70.
67. Ostergaard E, Rodenburg RJ, van den Brand M, *et al.* Respiratory chain complex I deficiency due to NDUF12 mutations as a new cause of Leigh syndrome. *Journal of medical genetics* 2011;48:737-40.
68. Ferreira M, Torracio A, Rizza T, *et al.* Progressive cavitating leukoencephalopathy associated with respiratory chain complex I deficiency and a novel mutation in NDUF1. *Neurogenetics* 2011;12:9-17.
69. Swalwell H, Kirby DM, Blakely EL, *et al.* Respiratory chain complex I deficiency caused by mitochondrial DNA mutations. *Eur J Hum Genet* 2011;19:769-75.
70. Scaglia F, Towbin JA, Craigen WJ, *et al.* Clinical spectrum, morbidity, and mortality in 113 paediatric patients with mitochondrial disease. *Paediatrics* 2004;114:925-31.
71. McFarland R, Taylor RW, Turnbull DM. A neurological perspective on mitochondrial disease. *Lancet Neurol* 2010;9:829-40.
72. Morava E, Rodenburg R, van Essen HZ, De Vries M, Smeitink J. Dietary intervention and oxidative phosphorylation capacity. *Journal of inherited metabolic disease* 2006;29:589.
73. Phoenix C, Schaefer AM, Elson JL, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord* 2006;16:814-20.







# Chapter 5

Clinical features and hetero-  
plasmy in blood, urine and  
saliva in 34 Dutch families  
carrying the m.3243A > G  
mutation



## Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A > G mutation

Paul de Laat<sup>1</sup>, **Saskia Koene**<sup>1</sup>, Lambert P.W.J. van den Heuvel<sup>1</sup>, Richard J.T. Rodenburg<sup>1</sup>, Mirian C.H. Janssen<sup>1</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc, Netherlands

Journal of Inherited Metabolic Disease 2012 Nov;35(6):1059-69

The m.3243A>G mutation has become known as the MELAS mutation. However, many other clinical phenotypes associated with this mutation have been described, the most frequently being Maternally Inherited Diabetes and Deafness. The m.3243A>G mutation can be detected in virtually all tissues, however heteroplasmy differs between samples. Recent reports indicate, a preference to perform mutation analysis in Urinary Epithelial Cells (UEC). To test this and to study the correlation between the mutational load in different tissues with two mitochondrial scoring systems (NMDAS and NPMDS), we investigated 34 families carrying the m.3243A>G mutation. Heteroplasmy was determined in three non-invasively collected samples, namely leucocytes, UEC and buccal mucosa. We included 127 subjects, of which 82 carried the m.3243A>G mutation. None of the children ( $n = 11$ ) had specific complaints. In adults ( $n = 71$ ), a median NMDAS score of 15 (IQR 10 - 24) was found. The most prevalent symptoms were hearing loss (68%), gastro-intestinal problems (59%), exercise intolerance (54%), and glucose intolerance (52%). Ten patients had neurologic involvement. Buccal mucosa had the best correlation with the NMDAS in all adults ( $r = 0.44$ ;  $p < 0.001$ ), whereas UEC had the strongest correlation with the NMDAS in severely affected patients ( $r = 0.59$ ;  $p = 0.002$ ). Heteroplasmy declined significantly with increasing age in all three samples ( $r(\text{leucocytes}) = -0.71$ ;  $p < 0.001$ );  $r(\text{UEC}) = -0.37$ ;  $p = 0.001$ );  $r(\text{buccal mucosa}) = -0.46$ ;  $p < 0.001$ ). In our cohort of 82 carriers, the m.3243A>G mutation causes a wide variety of signs and symptoms, MIDD being far more prevalent than MELAS. Looking at the characteristics of the three non-invasively available tissues for testing heteroplasmy we confirm that UEC are the preferred sample to test.

## Introduction

Mitochondria are responsible for production of adenosine triphosphate (ATP), through oxidative phosphorylation (OXPHOS). Mitochondrial dysfunction can result from mutations in either nuclear DNA or mitochondrial DNA (mtDNA). The incidence of congenital mitochondrial disorders based on an OXPHOS defect is at least 1:8,500 of all live births.<sup>1</sup> Mitochondrial DNA encodes 37 genes, 22 code for tRNAs, two for rRNAs and thirteen code for subunits of the OXPHOS complexes I, III, IV and V. Two important features of mitochondrial mutations are i) maternal inheritance, it can be assumed that all relatives from the maternal line are (dormant) carriers of the mutation. This applies only if there is not a *de novo* mutation. And ii) heteroplasmy, human cells contain over 100 – 1,000 mitochondria each, every mitochondrion contains one to ten copies of mtDNA. When all mtDNA copies are mutated, there is a homoplasmic mutation. When however not all mtDNA copies are mutated, there is a variation in mutation load and the mutation is called heteroplasmic.<sup>2</sup> Heteroplasmy can vary between different tissues of one patient and between one tissue in one patient in time.<sup>3</sup> This variation in heteroplasmy between tissues complicates the demonstration of a relationship between the severity of the disease in a non-accessible organ (e.g. the heart or the brain) and the degree of heteroplasmy in blood, urine or buccal mucosa.

The acronym MELAS was first used in 1984 by Pavlakis *et al.* to describe a group of patients with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes.<sup>4</sup> In 1990 the Adenine to Guanine transition at position 3243 of mitochondrial DNA (m.3243A>G) in the *MT-TL1* gene encoding tRNA<sup>LEU(UUR)</sup> was found as molecular basis for this disease.<sup>5,6</sup> A mutation in tRNA commonly causes a combined defect of the mtDNA encoded OXPHOS complexes. tRNA<sup>LEU(UUR)</sup> is partially (~12%) responsible for the incorporation of leucine into mitochondrial DNA encoded proteins, tRNA<sup>LEU(CUN)</sup> is responsible for the other part (~88%).<sup>7</sup> It is proposed that because of the minimal role of tRNA<sup>LEU(UUR)</sup> in incorporation of leucine, a mutation the tRNA<sup>LEU(UUR)</sup> is less lethal and therefore more frequently detected.<sup>7</sup>

The m.3243A>G mutation in the *MT-TL1* gene is the most common cause of MELAS syndrome, therefore the mutation is also known as the MELAS mutation. Later, more and more phenotypic expressions of the m.3243A>G mutation were found, including maternally inherited diabetes deafness (MIDD),<sup>8</sup> hypertrophic cardiomyopathy,<sup>9</sup> macular dystrophy,<sup>10</sup> focal segmental glomerulosclerosis (FSGS),<sup>11</sup> myoclonic epilepsy with ragged-red fibers (MERRF syndrome),<sup>12</sup> and oligosymptomatic variants of the acronym MELAS.<sup>13</sup>

Epidemiologic studies show a m.3243A>G mutation prevalence of 7.59/100,000 in the population of North-East England,<sup>14</sup> 16.3/100,000 in the population of Northern Finland<sup>15</sup> and up to 236/100,000 in the population of Australia.<sup>16</sup> These studies did not give a complete insight in the phenotypic expression of the m.3243A>G mutation in the subjects examined. The high mutation prevalence in the most recent study indicates a high number of undiagnosed carriers of the m.3243A>G mutation.

Since mitochondria and mtDNA are present in almost all tissues, heteroplasmy can theoretically be assessed in virtually every tissue. Two issues arise when testing heteroplasmy: the convenience of obtaining the sample and the differences in heteroplasmy levels between samples. For example, invasively obtained muscle tissue heteroplasmy usually gives a higher and more consistent level than conveniently attained blood, in the last case possibly leading to false-negative results. Recent studies showed a superiority of urine over blood as preferred non-invasive tissue for mutation analysis.<sup>3,17,18</sup> The relationship between mutation load and clinical phenotypes has been a subject of research for many years.<sup>19</sup> Whittaker *et al.* and Ma *et al.* recently showed a relationship between heteroplasmy levels in urinary epithelial cells and clinical symptoms in a small number of patients.<sup>17,20</sup>

Some of the many questions regarding the m.3243A>G mutation that we have include: Are all mutation carriers symptomatic? Should these carriers undergo screening for frequent and preventable symptoms such as cardiomyopathy, glucose intolerance, nephropathy or macular dystrophy? How should female carriers be counseled regarding fertility questions?

In this study, we clinically evaluated 34 Dutch families, probands and maternal relatives, carrying the m.3243A>G mutation. We provide data about the whole phenotypic spectrum of m.3243A>G mutation as well as information regarding the correlation between the level of heteroplasmy in different samples and the clinical severity of the disease.

## Patients and methods

### Patients

All probands are patients of the Nijmegen Centre for Mitochondrial Disorders at the Radboudumc, Nijmegen, The Netherlands, diagnosed with the m.3243A>G mutation in muscle or blood. All patients and maternal relatives were recruited by a letter, in which they were invited to participate in this study. This study was approved by the ethics committee of the Nijmegen-Arnhem region, The Netherlands. Written informed consent according to the Helsinki agreement was obtained from all parents and patients  $\geq 12$  years.

## Newcastle Mitochondrial Disease Scales

All patients were invited for a single visit to our outpatient clinic. Adult patients (>18yr) were scored using the Newcastle Mitochondrial Disease Adult Scale (NMDAS),<sup>21</sup> paediatric patients (≥18yr) were scored using the Newcastle Paediatric Mitochondrial Disease Scale (NPMDS).<sup>21</sup> The NMDAS and NPMDS constitute a validated method to monitor the clinical expression of mitochondrial disease and to follow-up the course of disease in time. The NMDAS and NPMDS consist of the following four sections. Section 1: Current function, gives insight into the general functioning of patients in past four weeks. Section 2: System-specific involvement, uses a clinical history supplemented by specific information to gain insight in the functioning of individual organ-systems. Section 3: Current clinical assessment, a general and neurological clinical examination, gives insight in the current functional status of the patient. This section includes three cognition tests. We were able to use the English symbol test, since it does not use language. For the reading test we used a Dutch equivalent test. In absence of a Dutch equivalent test for the Speed of Comprehension we could not use this test. We scored the cognition based on the symbol test and reading test only.

Section 1 to 3 of the NMDAS consists respectively of ten, nine and ten questions, which can be scored from 0 (no involvement) to 5 (severe involvement). For section 4, Quality of life (QoL), we used a Dutch translation of the SF-12v2 quality of life test. A score from 0 to 70 for mental and physical health, where 50 is the population mean, is obtained from this QoL score. The NPMDS exist respectively of seven, ten and nine questions, which can be scored from 0 (no involvement) to 3 (severe involvement). Section 4: Quality of life, uses a specially designed QoL questionnaire for children. These do not have a reference score, so we can only use them to show a difference in time, therefore they are not mentioned in this report. Intra-observer and inter-observer variability was shown to be low in both the NMDAS and the NPMDS.<sup>21</sup> All patients were scored independently by the same two investigators (PdL, SK), after which consensus was reached. The consensus score is used in the results of the study.

## m.3243A>G mutation analysis

At the same visit during which the NMDAS or NPMDS was scored, blood, saliva and urine was collected for heteroplasmy analysis. DNA was isolated from peripheral blood leukocytes using a salting-out method. The urine sediments and buccal swab samples were centrifuged for 10 minutes at 3000 rounds per minute, the pellet was washed with phosphate-buffered saline. DNA was extracted using a commercially available DNA isolation kit (Puregene™ DNA isolation kit; Gentra Systems, MN). DNA samples were analysed quantitatively using Pyrosequencing™ technology (Pyrosequencing, Uppsala, Sweden). Pyrosequencing was performed according to the protocol of the manufacturer. PCR of a mtDNA fragment containing the 3243 position was performed using the following primers: universal primer (biotinylated), 50-GGGACACCGCTGATC-

GTTTA-30; forward primer, 50-GACGGGACACCGCTGATCGTTTACAACCTTAGTA TTATACCCACAC-30; and reverse primer, 50-ATTAGAATGGGTACAATGAGGA-30. PCR was carried out in a 50 ml volume containing 0.02 mM forward primer, 0.2 mM reverse primer and 0.2 mM of the biotinylated universal primer. PCR conditions were 92°C for 30 s, 55°C for 30 seconds, 72°C for 30 seconds, for a total of 40 cycles. Single-stranded template DNA, which in the present assay is the forward strand of the fragment, was purified using streptavidin-coated Sepharose beads. The actual pyrosequencing was performed on the PSQ96 platform using sequence primer 50-TATGC-GATTACCGGGC-30. In a pyrosequence reaction, the four different deoxynucleotide triphosphates (dNTPs) are added separately one after the other. The incorporation of dNTP is accompanied by release of pyrophosphate (PPi). This PPi is involved in a light-producing reaction of which the amount of light produced is proportional to the number of nucleotides incorporated. The light is detected by a charge coupled device camera and seen as a peak in a pyrogram. Apyrase, a nucleotide-degrading enzyme, continuously degrades ATP involved in the light-producing reaction, and unincorporates dNTPs. This switches off light production and regenerates the reaction solution. Because the forward strand is used as template in the sequencing reaction, the change detected in the present assay concerns a T to C exchange. In fact, the amount of dTTP and dCTP incorporated at position 3243 during the sequencing reaction was determined in this way and from this the percentage of heteroplasmy was calculated. The pyrosequence reaction of the m.3234A>G mutation has a precision of 1.5%. The mutation is detected from a heteroplasmy level of 5%. The detection limit for the MELAS mutation (m.3243A>G) was determined by serial dilution of a sample containing this mutation with wild type mtDNA.

## Statistics

We used descriptive statistics to present patient characteristics and the results of the NMDAS and NPMDS. All data are presented as a median with interquartile ranges (IQRs). Non-parametric tests were used if the data did not reflect a Gaussian distribution. Spearman's rho correlation coefficient was used to evaluate the relationship between the clinical scores with the heteroplasmy in the different tissues. Pearson's correlation was used to correlate heteroplasmy in different levels to each other and to age. A *t*-test was used to compare the quality of life of the m.3243A>G mutation carriers to reference values.



## Results

### Patient characteristics

Hunderd-twenty-seven individuals from 34 families were included in the study. The cohort consists of 24 probands. Forty-eight family members were related in the first degree to a proband, 28 in the second degree, 18 in the third degree and nine in the fourth degree. The age of the patients ranged from two months to 80 years, median 39 years (IQR 25 – 54 years). Forty-two patients were male and 85 female (ratio 1:2.02). In 82 individuals the m.3243A>G mutation was detected in at least one of the samples. In 45 family members the mutation was not detected.

### Children

Eleven children (4 male, 7 female) carrying the m.3243A>G mutation were scored using the NPMDS. Two boys were probands, the other nine patients were maternal family members. The median age was 6 years (IQR 5.5 – 8.5 years). Five children scored zero points on the NPMDS, two children scored one point, three children scored two points and one child scored four points. The main complaints were a delay in development (in five children) in early childhood, which had been caught up by the time of investigation. One girl attended special school for learning difficulties. One boy had communicational problems both in his native language (Turkish) as in Dutch. Heteroplasmy in UEC in these children ranged from 8 - 93% and had a median level of 59% (IQR 32 – 77%).

### Adults

Seventy-one adult carriers (25 male, 46 female) of the m.3243A>G mutation were scored using the NMDAS. The median age was 45 years (IQR 34 – 55 years). They had a median clinical score of 4 (IQR 1 - 7) on Section 1; 4 (IQR 1 - 8) on section 2; and 3 (IQR 0 - 6) on Section 3. The total median score of Section 1-3 was 15 (IQR 10 - 24). Table 1 presents the symptoms sorted by prevalence. A complete overview of all results can be found in Figures 1A-C.

**Table 1.**

**Signs and Symptoms of the m.3243A>G mutation scored by the NMDAS.** Signs and symptoms scored by the NMDAS, sorted by prevalence ( $n = 71$ ).

Signs & Symptoms	Prevalence (%)
Hearing loss	68
Gastro-intestinal Symptoms	59
Decreased Vision	59
Exercise Intolerance	54
Glucose Intolerance	52
Gait Instability	51

Cerebellar Ataxia	49
Myopathy	48
Cognition < 33th percentile	45
Ptosis	45
Cardiovascular Involvement	44
Vision Complaints	37
Respiratory Muscle Weakness	37
Problems Dressing	34
Problems Hygiene	34
Speech Difficulties	31
Psychiatric Problems	28
Migraine Headaches	25
Neuropathy	20
Swallowing Difficulties	18
Dysphonia / Dysarthria	17
Apraxia	13
Handwriting Difficulties	13
Seizures	13
CPEO	7
Encephalopathic Episodes	6
Stroke-like Episodes	6
Pyramidal Involvement	6
Extrapyramidal Involvement	3

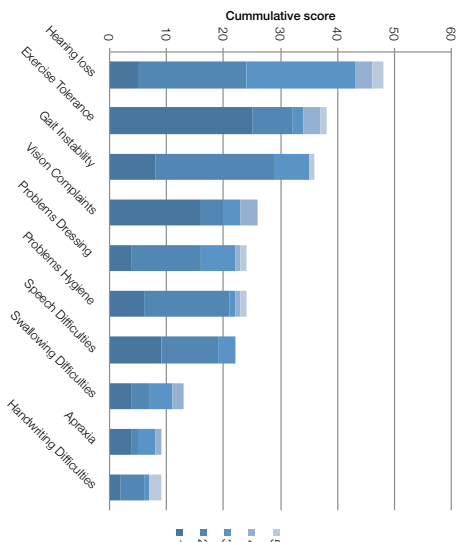
*CPEO = Chronic progressive ophthalmoplegia*

### Summary of the most important findings:

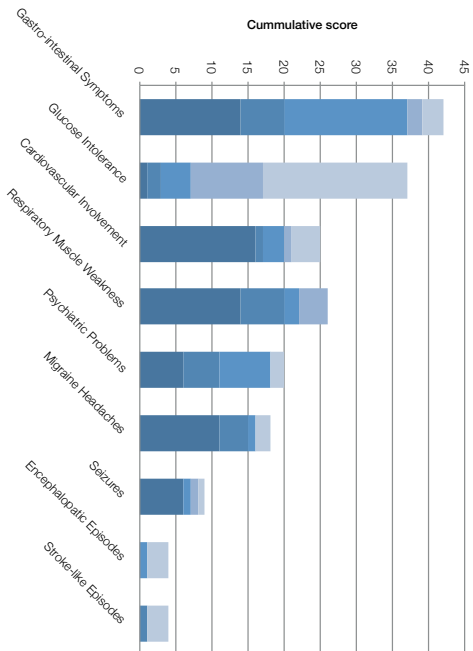
Hearing: 48 carriers (67.6%) had hearing difficulties; 24 carriers experienced mild deafness; whereas another 19 carriers had moderate deafness (not fully corrected with hearing aid); three carriers had severe deafness (poor even with hearing aid) and two carriers had end stage deafness requiring a cochlear implant.

Exercise Tolerance: 38 carriers (53.5%) had a compromised exercise tolerance; 25 were only limited on inclines or stairs; whereas the other carriers were also limited on flat services.

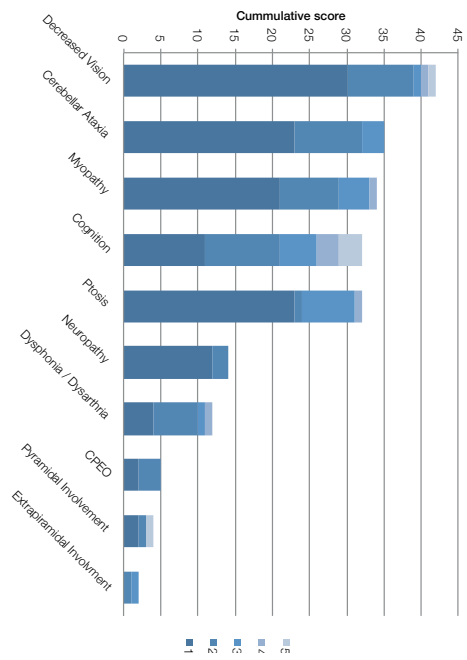
NMDAS Section 1



NMDAS Section 2



NMDAS Section 3



**Figure 1 A-C.**  
All clinical features scored by the NMDAS. Results of the Newcastle Mitochondrial Disease Adult Scale (NMDAS) are shown divided in the three sections of the scale, for all adult carriers of the m.3243A>G mutation ( $n = 71$ ). A) Section 1: Current function; B) Section 2: System-specific involvement; C) Section 3: Current clinical assessment.

Gait stability and cerebellar ataxia: 36 carriers (50.7%) reported difficulties maintaining their stability; 35 carriers (49.3%) had difficulties with the heel-toe test at the clinical assessment; eight carriers reported difficulties on uneven grounds; and 21 carriers reported occasional balance problems when walking. Six carriers reported occasional fall assessment 23 carriers had a hesitant heel-toe test; nine were unable to maintain heel-toe walking and three carriers were unable to walk heel-toe.

Gastro-intestinal Symptoms: 42 carriers (59.2%) reported gastro-intestinal symptoms. 14 had mild constipation; six had occasional symptoms of irritable bowel; 17 carriers had severe constipation (requiring daily medication) or severe irritable bowel complaints; in five carriers gastro-intestinal symptoms were that severe they needed hospital admission; three of them underwent surgical procedures for gastro-intestinal dysmotility.

Diabetes Mellitus: 37 carriers (52.1%) had a compromised glucose tolerance, for which ten were treated with tablets (non insulin depending diabetes mellitus) and 20 required insulin treatment (insulin depending diabetes mellitus); four patient were on a diet and two carriers received no treatment for impaired glucose tolerance, one patient reported diabetes gravidarum.

Ptosis: 32 carriers (45.1%) had a ptosis, mild ptosis (not obscuring either pupil) occurred in 23 carriers; moderate ptosis occurred in eight carriers; whereas one patient had a severe ptosis (obscuring bilateral >1/3 of pupils).

Myopathy: 34 carriers (47.9%) had reduced muscle strength; 21 carriers had minimal muscle weakness in hip flexion or shoulder abduction only (MRC 4+/5); eight carriers had mild proximal muscle weakness (MRC 4/5) and four carriers had moderate muscle weakness (MRC 4) with difficulty to rise from 90 degree squat; one patient was unable to rise from 90 degree squat.

Cardiomyopathy: 25 carriers (38.5%) had cardiac involvement; 16 had asymptomatic changes on electrocardiogram, mainly repolarisation disturbances; one had asymptomatic left ventricular hypertrophy; seven had cardiomyopathy of which four had a ejection fraction of less than 30%; one patient had a pacemaker for a complete AV-block.

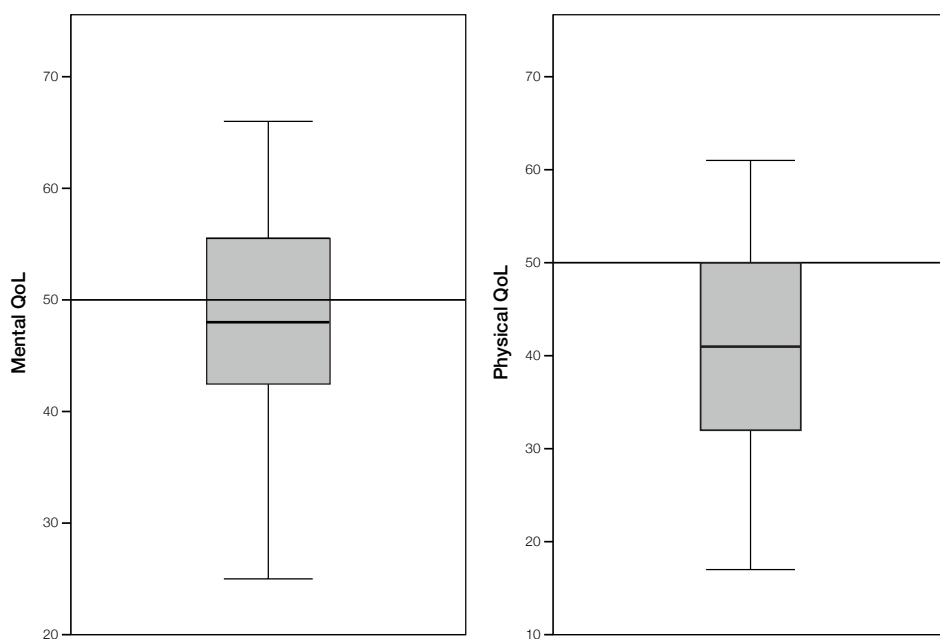
Epilepsy: A history of seizures was present in nine carriers, of which three had had seizures in thalls because of balance problems; one patient was unable to walk unsupported onclinic past year.

Stroke-like episodes: A history of stroke-like episodes was present in four carriers, of which three had stroke-like episodes in the past year.

Further neurologic examination: Twelve carriers had subtle sensory symptoms or areflexia (in knee and elbow reflexes); four carriers showed symptoms of pyramidal involvement and two carriers showed symptoms of extrapyramidal involvement.

Section 4 of the NMDAS scores QoL. It distinguishes Quality of Physical Health (QoL<sub>P</sub>) and Quality of Mental Health (QoL<sub>M</sub>). QoL<sub>P</sub> had a range of 17 to 61, with a median of 41 (IQR 32 - 50). QoL<sub>M</sub> had a range of 25 to 66, with a median of 48 (IQR 42 - 48; Figure 2).

Using a *t*-test a significantly ( $p < 0.001$ ) lower QoL<sub>P</sub> was found in patients carrying the m.3243A>G mutation against the test average of 50. No difference in QoL<sub>M</sub> in carriers against the test average could be found ( $p = 0.055$ ). Comparing QoL in carriers to non-carriers, there was a significant lower QoL<sub>P</sub> in carriers ( $p < 0.001$ ) and no difference in QoL<sub>M</sub> ( $p = 0.40$ ). The QoL<sub>P</sub> was significantly negatively correlated to the NMDAS ( $r = -0.57$ ;  $p > 0.001$ ). There was no correlation between QoL and the heteroplasmy in any of the samples.



**Figure 2 .**

**Quality of Life. Section 4 of the NMDAS is a QoL test, divided in QoL<sub>P</sub> and QoL<sub>M</sub>.** All adult carriers of the m.3243A>G mutation are shown ( $n = 71$ ). The line at 50 shows the US average of the test. QoL<sub>P</sub> is significantly lower than the US average.

## Heteroplasmy Levels

Heteroplasmy levels in leucocytes were determined in 126 individuals (one child refused the blood draw), in UEC in 121 individuals (insufficient material was received at the laboratory in six individuals) and in buccal mucosa in all 127 individuals. Eighty-two individuals had a heteroplasmy level of  $\geq 5\%$  in at least one of the samples. In the carriers, the mean heteroplasmy in leucocytes was 22% (range 2 - 65%); in UEC 48% (range 4 - 96%) and 35% (range 2 - 74%) in buccal mucosa. Despite these differences between the samples there were strong pair wise correlations between the samples ( $r = 0.66 - 0.85$ ;  $p < 0.001$ ; Figures 3A-C). A negative correlation between the level of heteroplasmy and age was present in all samples, for leucocytes ( $r = -0.68$ ;  $p < 0.001$ ), for UEC ( $r = -0.34$ ;  $p = 0.003$ ), and for buccal mucosa ( $r = -0.43$ ;  $p < 0.001$ ), Figures 4A-C).

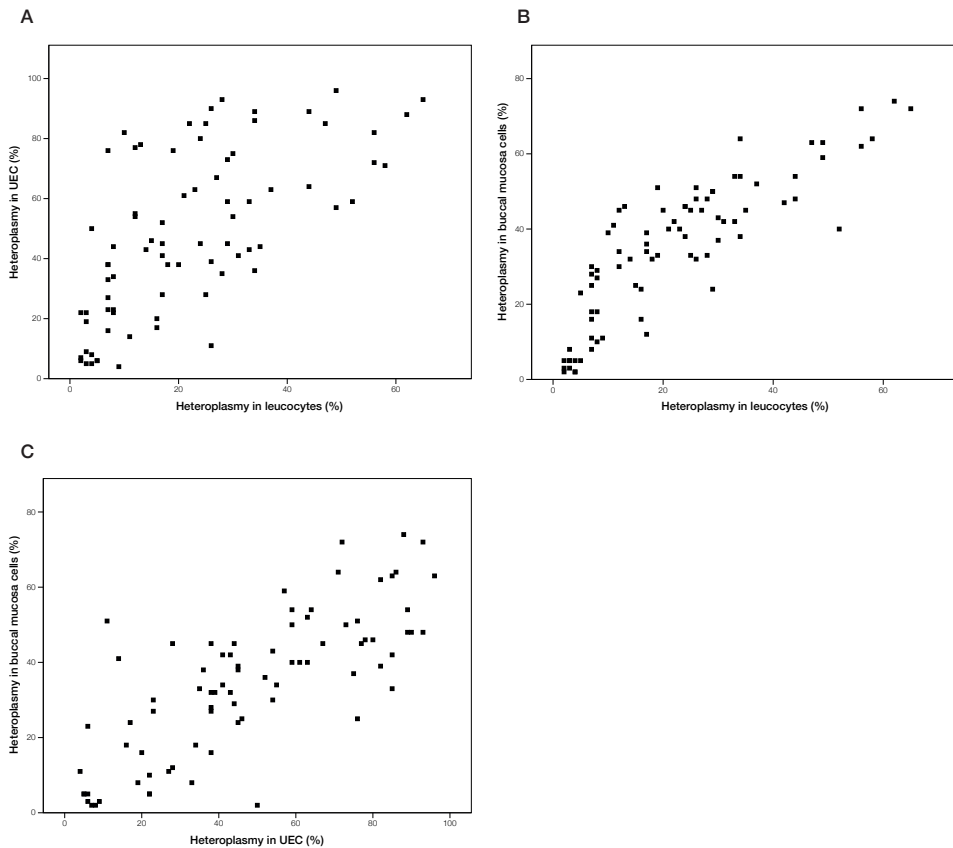
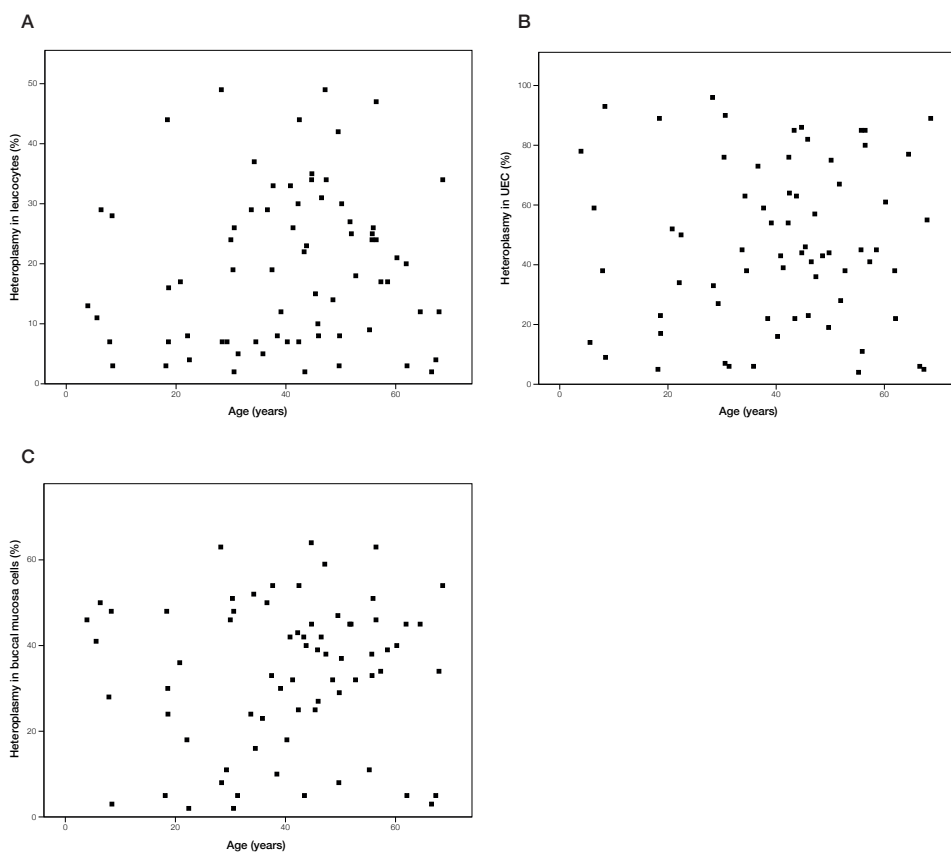


Figure 3A-C.

**Heteroplasmy levels in all patients carrying the m.3243A>G mutation.** Heteroplasmy levels in all patients carrying the m.3243A>G mutation ( $n = 82$ ) in three different samples (leucocytes, UEC and buccal mucosa) are correlated.

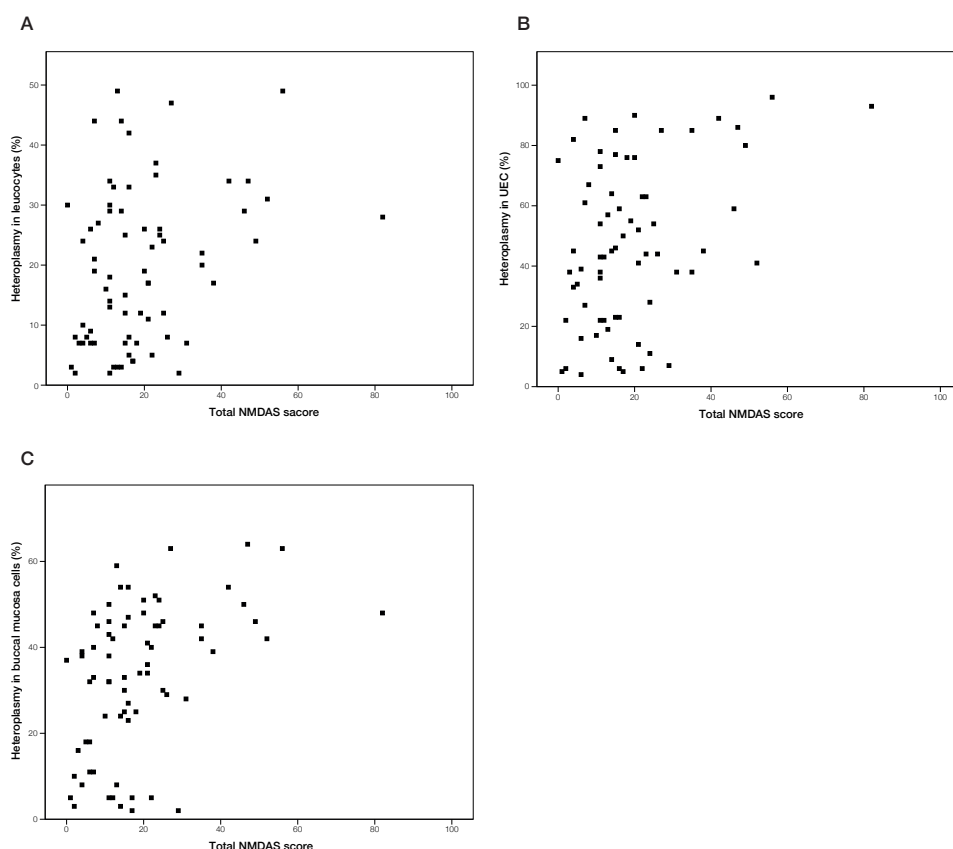
The heteroplasmy level in UEC was below the detection limit in one (1%) carrier that did have a detectable mutation load in buccal mucosa and leucocytes. Heteroplasmy levels in buccal mucosa were below the detection limit in seven (9%) carriers that did have a detectable mutation load in UEC. Heteroplasmy levels in leucocytes were below the detection limit in 12 (15%) carriers that did have a detectable mutation load in UEC.



**Figure 4A-C.**

**Heteroplasmy versus age.** Heteroplasmy levels in all patients carrying the m.3243A>G mutation ( $n = 82$ ) show a negative correlation to age in leucocytes, UEC and buccal saliva.

A correlation between clinical condition of the patient and the level of heteroplasmy could not be made using the NPMDs in children, because of the limited number of paediatric carriers included in the study. In adults, there was a correlation between the score on the NMDAS and the heteroplasmy, in leucocytes ( $r = 0.25$ ;  $p = 0.032$ ), in UEC ( $r = 0.29$ ;  $p = 0.016$ ), and in buccal mucosa ( $r = 0.43$ ;  $p < 0.001$ ; Figures 5A-C). Patients with a NMDAS of more than 20 ( $n = 25$ ), have a better correlation between the clinical scores and the heteroplasmy levels in leucocytes ( $r = 0.37$ ;  $p = 0.072$ ), and UEC ( $r = 0.53$ ;  $p = 0.008$ ).



**Figure 5A-C.**

**Heteroplasmy vs NMDAS score.** Heteroplasmy levels in leucocytes, UEC and buccal saliva in all adult patients carrying the m.3243A>G mutation ( $n = 71$ ) show a correlation to the score on the NMDAS.



## Discussion

Following the results of this study, we propose to introduce a new term in the mitochondrial genetic nomenclature: dormant carrier. The term 'dormant carrier' is used to indicate that a patient has a mitochondrial mutation, but has no obvious clinical symptoms of the mutation (yet). We have chosen this new term because we experienced that the nomenclature in mitochondrial inheritance has an overlap with nomenclature in Mendelian inheritance, which causes confusion in providing patient information. In the Mendelian nomenclature, a carrier is a patient that has a recessive mutation on one allele, and is never going to be a patient. On the contrary, a dormant carrier of a mitochondrial mutation may at some point become symptomatic (awake), depending on heteroplasmy levels and unknown other factors. In patient counseling and preventive medicine this makes a big difference.

Since the Newcastle scores consist of subjective questions regarding the personal opinion of the patient on daily functioning (Section 1) and the interpretation of the findings at physical examination (Section 3), all patients were scored separately by two investigators. Only ten patients aged under eighteen were included in the study. All ten children had no or only minor complaints. However, we did include four families in which a child had died earlier because of symptoms of the m.3243A>G mutation, so no definite assumption can be made concerning the phenotype of the m.3243A>G mutation in children based on this study. The study does however show that children can be free of symptoms even with a high level of heteroplasmy (up to 93% in UEC). The adult patients with high levels of heteroplasmy developed complaints in late infancy to early adolescence. The younger children with high levels of heteroplasmy should therefore be closely monitored to identify early symptoms of the clinical expression of the mutation at an early stage.

The m.3243A>G mutation is most known for its relation with MELAS syndrome. In this study only one patient had stroke-like episodes in the past year and three other patients had stroke-like episodes in the past. Other neurologic symptoms as seizures ( $n = 9$ ) and encephalopathy ( $n = 4$ ) occurred also infrequent. Giving an indication that MELAS syndrome is not the primary expression of the m.3243A>G mutation in this cohort. This finding is in line with recent reports<sup>16</sup> that the m.3243A>G mutation is much more prevalent than previously thought. Patients presenting with symptoms of MELAS syndrome are often referred to a specialised centre in which exploratory research for mitochondrial mutation is performed. A patient with the m.3243A>G mutation with a MELAS syndrome phenotype is therefore unlikely to be missed.

Forty-eight patients (67.6%) of this study had self-reported hearing loss. Thirty-seven patients (52.1%) have diabetes mellitus. Only six of the 37 patients with diabetes

mellitus do not have hearing loss. Making a total of 31 patients in this study that have diabetes deafness, giving an indication that MIDD is a more frequent expression of the m.3243A>G mutation than MELAS syndrome. The combination of diabetes mellitus and hearing loss is however less a trigger to refer a patient to a specialised centre. So a mutation analysis is less frequently performed.

Besides the typical symptoms of MELAS syndrome or MIDD we observed high frequencies of other symptoms. Cardiac involvement, gastro-intestinal symptoms and visual disturbances are most frequent. Cardiac involvement has previously been described.<sup>9,22,23</sup> In our cohort we found a large group of patients with asymptomatic ECG changes on one hand, but some severely affected patients as well. In the four patients with an ejection fraction of less than 30% the cardiac problems dominate the clinical picture. A large number of patients in our cohort exhibits gastro-intestinal problems. In most cases patients suffer from mild to moderate constipation or symptoms of irritable bowel syndrome. In five patients symptoms required admission or even surgical intervention. Cases of chronic intestinal pseudo obstruction have previously been described in patients with the m.3243A>G mutation.<sup>24,25</sup> Twenty-six percent of the patients reported visual disturbances and 42% of the patients did not have a vision of 1.0 on a brief visual acuity test. Previous reports indicate the presence of macular dystrophy in patients with the m.3243A>G mutation.<sup>10,26</sup> Because of this multi-organ involvement of the m.3243A>G mutation we recommend that all carriers of the mutation are checked for early stages of these symptoms by specialised physicians, including a consultation by a cardiologist and an ophthalmologist, also when there is no diabetes present.

We assessed quality of life in our cohort using the SF-12v2 questionnaire. A lower QoL<sub>P</sub> is negatively correlated to the NMDAS. Indicating, as suspected, that when a patient has more symptoms of the m.3243A>G mutation, a lower QoL<sub>P</sub> is reported. QoL<sub>M</sub> was not different from healthy controls nor from their relatives that do not carry the mutation. Twenty patients did report a history of depression or anxiety. This was however not significantly different from their relatives that did not carry the mutation.

As indicated in Figure 5A-C we found a correlation between NMDAS and heteroplasmy in all three samples, of which buccal mucosa had the best correlation ( $r = 0.43$ ;  $p < 0.001$  versus  $r = 0.25$ ;  $p = 0.032$  in leucocytes and  $r = 0.29$ ;  $p = 0.016$  in UEC). When we selected for the most severely affected patients (NMDAS >20) we found stronger correlation for leucocytes ( $r = 0.37$ ;  $p = 0.072$ ) and UEC ( $r = 0.53$ ;  $p = 0.008$ ). Whittaker *et al.* 20 was the first to describe a correlation between heteroplasmy and symptoms. We confirm the good ability of UEC heteroplasmy to predict clinical outcome but add that this accounts most for the more severely affected patients. In the general group of carriers of the m.3243A>G mutation heteroplasmy in buccal saliva has a better ability to predict clinical outcome.

Different studies have described a decrease in heteroplasmy with increasing age.<sup>3,13,27</sup> As demonstrated in Figure 4A-C, we also find a decreasing heteroplasmy in all samples. The decrease in heteroplasmy with increasing age is much more prominent in leucocytes ( $r = -0.71$ ;  $p < 0.001$ ) than in the other samples ( $r(\text{UEC}) = -0.37$ ;  $p = 0.001$ ;  $r(\text{buccal mucosa}) = -0.46$ ;  $p < 0.001$ ). However the decrease in heteroplasmy level with age can only be confirmed after a prospective follow-up study. We also found that heteroplasmy in UEC has a 26 percent point higher value than heteroplasmy in leucocytes and a 11 percent point higher value than heteroplasmy in buccal mucosa. The m.3243A>G mutation was undetectable in leucocytes in 12 patients in which the mutation was detected in UEC. Of these 12 patients the mutation was also undetectable in buccal mucosa in seven patients. There was one patient in which the m.3243A>G mutation could not be detected in UEC, where it was detected in leucocytes and buccal mucosa. However we should take in account that we can never predict the values of untested tissues as muscle, heart, kidney, liver or brain. This makes that UEC are the best non-invasively available tissue to test if a patient is suspected of the m.3243A>G mutation.

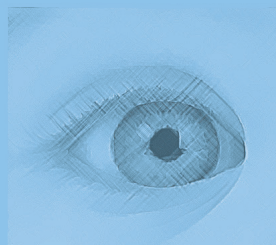
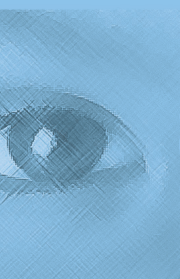
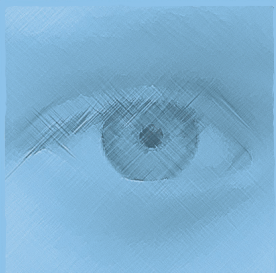
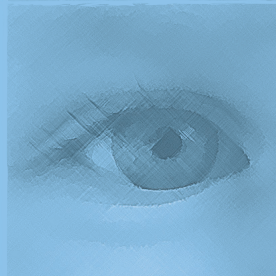
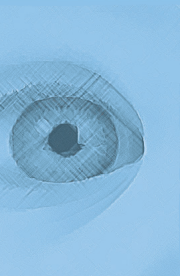
## Conclusion

We conclude that the m.3243A>G mutation causes a wide variety of signs and symptoms, MIDD being the most prevalent phenotypic expression. Of the three non-invasively available tissues for testing heteroplasmy, we advice to use UEC to detect the mutation. Dormant carriers should be checked regularly by a specialised physician to see whether the disease has awoken and to be early in preventing and treating symptoms.

## References

1. Chinnery PF, Turnbull DM. Epidemiology and treatment of mitochondrial disorders. *Am J Med Genet* 2001;106:94-101.
2. Cree LM, Samuels DC, Chinnery PF. The inheritance of pathogenic mitochondrial DNA mutations. *Biochimica et biophysica acta* 2009;1792:1097-102.
3. Frederiksen AL, Andersen PH, Kyvik KO, Jeppesen TD, Vissing J, Schwartz M. Tissue specific distribution of the 3243A->G mtDNA mutation. *Journal of medical genetics* 2006;43:671-7.
4. Pavlakis SG, Phillips PC, DiMauro S, De Vivo DC, Rowland LP. Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: a distinctive clinical syndrome. *Annals of neurology* 1984;16:481-8.
5. Goto Y, Nonaka I, Horai S. A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* 1990;348:651-3.
6. Kobayashi Y, Momoi MY, Tominaga K, *et al.* A point mutation in the mitochondrial tRNA(Leu)(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). *Biochemical and biophysical research communications* 1990;173:816-22.
7. Wittenhagen LM, Kelley SO. Impact of disease-related mitochondrial mutations on tRNA structure and function. *Trends in Biochemical Sciences* 2003;28:605-11.
8. van den Ouweland JM, Lemkes HH, Ruitenbeek W, *et al.* Mutation in mitochondrial tRNA(Leu)(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nature genetics* 1992;1:368-71.
9. Lev D, Nissenkorn A, Leshinsky-Silver E, *et al.* Clinical presentations of mitochondrial cardiomyopathies. *Paediatric cardiology* 2004;25:443-50.
10. Michaelides M, Jenkins SA, Bamiou DE, *et al.* Macular dystrophy associated with the A3243G mitochondrial DNA mutation. Distinct retinal and associated features, disease variability, and characterization of asymptomatic family members. *Archives of ophthalmology* 2008;126:320-8.
11. Lowik MM, Hol FA, Steenbergen EJ, Wetzels JF, van den Heuvel LP. Mitochondrial tRNA<sup>Leu</sup>(UUR) mutation in a patient with steroid-resistant nephrotic syndrome and focal segmental glomerulosclerosis. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* 2005;20:336-41.
12. Durand-Dubief F, Ryvlin P, Manguiere F. [Polymorphism of epilepsy associated with the A3243G mutation of mitochondrial DNA (MELAS): reasons for delayed diagnosis]. *Revue neurologique* 2004;160:824-9.
13. Koopman WJ, Hink MA, Verkaart S, Visch HJ, Smeitink JA, Willems PH. Partial complex I inhibition decreases mitochondrial motility and increases matrix protein diffusion as revealed by fluorescence correlation spectroscopy. *Biochimica et biophysica acta* 2007;1767:940-7.
14. Chinnery PF, Johnson MA, Wardell TM, *et al.* The epidemiology of pathogenic mitochondrial DNA mutations. *Annals of neurology* 2000;48:188-93.
15. Majamaa K, Moilanen JS, Uimonen S, *et al.* Epidemiology of A3243G, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population. *American journal of human genetics* 1998;63:447-54.
16. Manwaring N, Jones MM, Wang JJ, *et al.* Population prevalence of the MELAS A3243G mutation. *Mitochondrion* 2007;7:230-3.
17. Ma Y, Fang F, Yang Y, *et al.* The study of mitochondrial A3243G mutation in different samples. *Mitochondrion* 2009;9:139-43.

18. Marotta R, Reardon K, McKelvie PA, *et al.* Association of the MELAS m.3243A>G mutation with myositis and the superiority of urine over muscle, blood and hair for mutation detection. *J Clin Neurosci* 2009;16:1223-5.
19. Chinnery PF, Howell N, Lightowlers RN, Turnbull DM. Molecular pathology of MELAS and MERRF. The relationship between mutation load and clinical phenotypes. *Brain* 1997;120 (Pt 10):1713-21.
20. Whittaker RG, Blackwood JK, Alston CL, *et al.* Urine heteroplasmy is the best predictor of clinical outcome in the m.3243A>G mtDNA mutation. *Neurology* 2009;72:568-9.
21. Phoenix C, Schaefer AM, Elson JL, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord* 2006;16:814-20.
22. Wortmann SB, Rodenburg RJ, Backx AP, Schmitt E, Smeitink JA, Morava E. Early cardiac involvement in children carrying the A3243G mtDNA mutation. *Acta Paediatr* 2007;96:450-1.
23. Fayssoil A. Heart diseases in mitochondrial encephalomyopathy, lactic acidosis, and stroke syndrome. *Congestive heart failure (Greenwich, Conn)* 2009;15:284-7.
24. Verny C, Amati-Bonneau P, Letournel F, *et al.* Mitochondrial DNA A3243G mutation involved in familial diabetes, chronic intestinal pseudo-obstruction and recurrent pancreatitis. *Diabetes & metabolism* 2008;34:620-6.
25. Bergamin CS, Rolim LC, Dib SA, Moises RS. Unusual occurrence of intestinal pseudo obstruction in a patient with maternally inherited diabetes and deafness (MIDD) and favorable outcome with coenzyme Q10. *Arquivos brasileiros de endocrinologia e metabologia* 2008;52:1345-9.
26. Massin P, Virally-Monod M, Vialettes B, *et al.* Prevalence of macular pattern dystrophy in maternally inherited diabetes and deafness. GEDIAM Group. *Ophthalmology* 1999;106:1821-7.
27. Rahman S, Poulton J, Marchington D, Suomalainen A. Decrease of 3243A>G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *Am J Hum Genet.* 2001 Jan;68(1):238-40.



# Part III

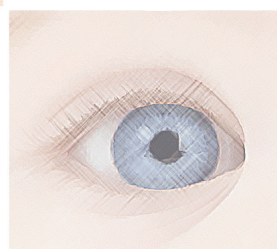
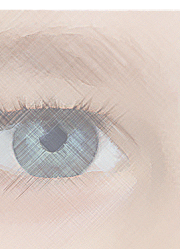
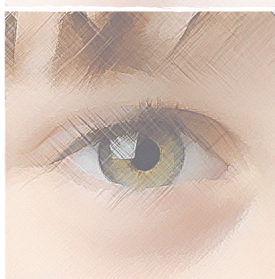
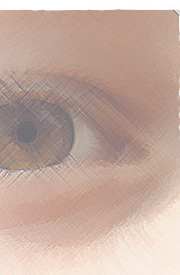
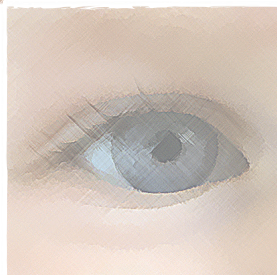
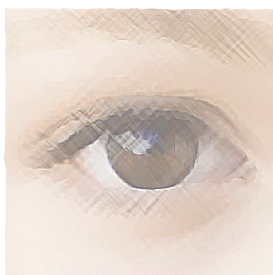
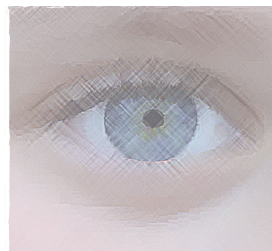
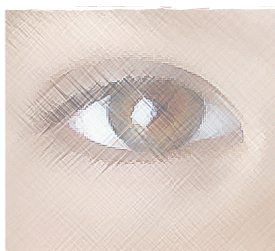
## Validation studies





## Goal

To obtain experience with the previously selected outcome measures in children with mitochondrial disorders.



# Chapter 6

The value of using 3D accelerometry in estimating daily physical activity in children with mitochondrial disease



## The value of using 3D accelerometry in estimating daily physical activity in children with mitochondrial disease

**Saskia Koene**<sup>1</sup>, Ilse Dirks<sup>1</sup>, Esmee van Mierlo<sup>1</sup>, Pascal de Vries<sup>1</sup>, Anjo J.W.M. Janssen<sup>2</sup>, Jan A.M. Smeitink<sup>1</sup>, Arjen Bergsma<sup>2,3</sup>, Hans Essers<sup>4</sup>, Kenneth Meijer<sup>4</sup>, Imelda J.M. de Groot<sup>1,3</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders (NCMD), Radboud university medical centre, Nijmegen, the Netherlands

<sup>2</sup>Department of Pediatric Rehabilitation, Pediatric Physical Therapy, Radboud university medical centre

<sup>3</sup>Radboud University Medical Centre, Donders Centre for Neuroscience, Department of Rehabilitation

<sup>4</sup>Department of Human Movement Sciences, Maastricht University Medical Centre +

Submitted

Lack of energy and tiredness are among the most burdensome complaints experienced by children with mitochondrial disease and therefore, outcome measures for these symptoms are urgently needed. Many laboratory tests and questionnaires to assess fatigability are not feasible in children with mitochondrial disease. Moreover, measuring the activity at home may better reflect the lack of energy and tiredness experienced by these children than a laboratory test. The aim of this study was to test the feasibility, validity, and test-retest reliability of measuring daily physical activity with 3D accelerometers placed on chest, upper arm, lower arm, leg and if available wheelchair in children with mitochondrial disease and compare them to healthy age- and gender-matched controls.

In 17 children with confirmed mitochondrial disease and 16 peers, we found that measuring daily physical activity during the whole weekend was practically feasible in all participants. Four percent of the measurements failed because of hardware problems, leading to the exclusion of one healthy child. We found good face validity by visually correlating the validation videos and activity diaries to the accelerometer data-graphs. Moreover, accelerometry confirmed lower dynamic activity (*i.e.* walking) in non-ambulatory patients compared to ambulatory patients. Patients with mitochondrial disorders had significantly lower peak-intensity and were resting more compared to their age- and gender-matched peers. Testing eight children in another season with significantly different temperature and hours of sun showed good test-retest reliability.

We conclude that accelerometry is a promising measurement tool for future studies in children with mitochondrial disease. Before using the accelerometer as an outcome measure in clinical trials, technical failures will have to be solved and family activities should be standardised.

## Introduction

Many potential therapies have been attempted in patients with mitochondrial disorders, however to date, none of them has proven beneficial.<sup>1</sup> This is partly due to a lack of sensitive and responsive outcome measures. In a recent review of all published studies in mitochondrial disease, an international expert panel recommended (amongst others) to use validated and clinically meaningful end points.<sup>1</sup> Only very few studies investigating outcome measures in paediatric mitochondrial disease have been published and most of the outcome measures studied are not generally applicable among children with mitochondrial disease.

Since lack of energy and fatigue are among the most burdensome complaints experienced by children with mitochondrial disease and their parents,<sup>2</sup> this symptom should be covered in future clinical trials. However, testing fatigue or fatigability in this paediatric population is challenging, since many children are not able to rate their fatigue using one of the widely used self-report fatigue questionnaires<sup>3,4</sup> because of intellectual disabilities and most endurance tests (e.g. cycle ergometry<sup>5</sup>) are not feasible or too burdensome for children with mitochondrial disorders. Besides, measuring performance in a laboratory situation may not always reflect the disabilities experienced in daily life since the performance in daily life may differ from the abilities of the child.<sup>6-8</sup>

We hypothesised that measuring daily physical activity at home is a clinically relevant outcome measure and a good reflection of the fatigue experienced by children with mitochondrial disease in daily life. Physical activity is defined as ‘any bodily movement produced by skeletal muscle contraction that results in caloric expenditure’ and includes sports, hobbies, playing, walking, cycling and activities of daily living.<sup>9</sup> Of the many dimensions of physical activity (type, intensity, frequency, and duration), we were most interested in the intensity, frequency and the duration of movement. We hypothesised that the physical activity of children with mitochondrial disease would be less intense and consisting of shorter active periods with longer periods of resting.

Daily physical activity in a home situation can be measured by using 3D accelerometry.<sup>10-16</sup> For this study, we selected the MOX-accelerometer, a device with opportunities to design a tailored analysis and to measure for long periods of time (14 days). Here, we aim to test the feasibility, validity and test-retest reliability of using 3D accelerometry as a predictor for daily activity of children with mitochondrial disease.

## Methods

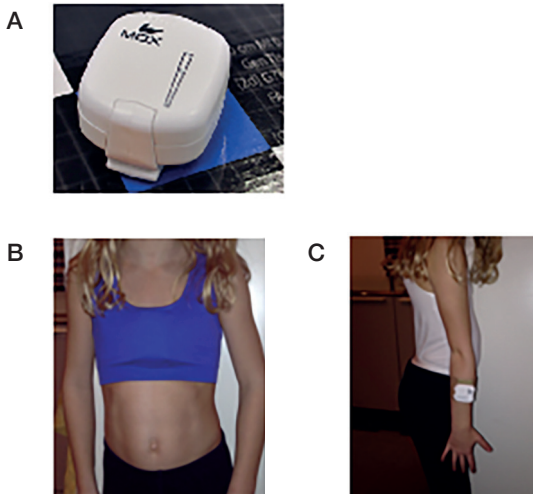
This is an observational study, testing the feasibility, validity and test-retest reliability of various parameters that were calculated based on accelerations measured by 3D accelerometers in indicating physical activity of children with mitochondrial disorders.

### MOX-accelerometer

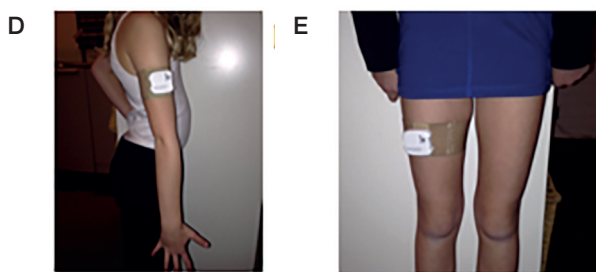
For this study, we used a MOX accelerometer (MOX sensor, model MMOXX1.01, Maastricht Instruments BV, The Netherlands), that measures accelerations (range  $\pm 6G$ ) in three degrees of freedom with a sample frequency of 25Hz. The acceleration data was filtered with a Butterworth 0.025-7.5Hz 4th-order high-pass filter to remove noise and movement artefacts.

A set of four or five sensors was used. The accelerometers were attached to the chest, dominant lower arm and upper arm and to the leg using an attachment band (limbs) or a top (chest; Figure 1). If the patient used a wheelchair, a fifth accelerometer was attached to the wheelchair. The wheelchair sensor was used to indicate passive moments of the child.

To estimate the amount of daily activity, various parameters were calculated from the acceleration data of each sensor, by using Matlab procedures that were developed before.<sup>17</sup> The first parameter are the activity counts, which was calculated by integrating the acceleration over 1-minute episodes and summing this outcome over all three axes. A constant acceleration of 1G (gravitational constant) over 1 minute corresponds with 1,000 counts.







**Figure 1.**

**The accelerometer localisations.** A) The MOX-accelerometer is approximately 4.5 x 4.0 x 1.4 cm in size and weighs 27 grams; B) Attachment of the accelerometer to the (dominant) upper leg; C) to the chest; D) the upper- and E) the lower arm.

We used the following outcome measures: i) average amount of counts per hour the sensors were worn (average counts (total amount of counts measured with the sensor/worn hours; counts/hour), also referred to activity level); ii) the maximal intensity (maximal amount of counts per min (counts/min)); and iii) the largest area under the curve (AUC) during 30 minutes (largest AUC during ½ hour (counts)). The second outcome measure is an activity classification which categorises the performed activities per second into lying, standing or being dynamically active. Lying and standing are classified depending on the gravitational angle acting on the posterior-anterior and cranial-caudal axes. Being dynamically active is classified when the integration of the acceleration over 1-second episodes is above a pre-defined threshold.

## Study protocol

### Study protocol for patients

Patients were recruited at the Nijmegen Centre for Mitochondrial Disorders. Patients aged 4 - 18 years old with a confirmed mitochondrial disease, either based on pathological mutations in mtDNA or nuclear DNA or on mitochondrial dysfunction in fresh muscle as measured by routine biochemistry as applied in our centre, were eligible for inclusion. Exclusion criteria: i) expected by the treating physician that travelling to the hospital would be too burdensome to the patient; ii) fever; iii) epilepsia continua; or iv) altered state of consciousness compared to normal at the time of inclusion.

Patients were assessed at the outpatient clinic of the Nijmegen Centre for Mitochondrial Disorders (NCMD) on Fridays. An experienced paediatric physiotherapist performed the Gross Motor Function Measure-88 (GMFM) and the Modified Tardieu test for spasticity to assess the motor abilities and the presence of spasticity in each individual child.

The GMFM is an instrument measuring gross motor abilities in children. Though it was originally designed for and validated in patients with cerebral palsy,<sup>18-20</sup> the GMFM is able to measure reliably in neuromuscular disorders as well.<sup>21</sup> The GMFM is expressed as a percentage of the highest score. The modified Tardieu test measures spasticity by passively moving the joint at specified velocities and scoring the angle at which the muscle reaction to stretch is felt.<sup>22, 23</sup> We performed the Modified Tardieu of the ankles (mm. gastrocnemicus), the elbows (mm. biceps brachii) and the knees (hamstrings) bilaterally in patients with increased muscle tone.

After the two tests, parents were instructed how to attach the accelerometers. Wearing the sensors, patients were – if possible - instructed to follow a validation protocol (standardized activities, including waving, throwing a ball, lying down, sitting, standing, walking and running). In case of limited physical abilities, the position (orientation) of the arm, leg and chest was changed passively, if possible at low, middle and high velocity (intensity). The patient was videotaped with a synchronized camera during all tests, to be able to correlate specific movements (e.g. raising an arm, walking, movement disorders or epilepsy) to the data obtained by the accelerometer. By correlating these video images with the accelerometer data-graphs (correspondence of orientation and intensity for each sensor) we determined the face validity of the measurements in a laboratory situation.

After completion of the validation protocol, patients were asked to wear the sensors over the weekend, while the parents completed a diary, rating the happiness and the activity of the child. The accelerometry measurements were located at the patients' home environments. Parents were asked to complete the diary with the exact timing and a description of the activity (e.g. 12:36 – 13:18: Lunch, independently eating bread with knife and fork). During the weekend, patients were asked to wear the accelerometers at all waking hours, with the exception of bathing, showering and swimming. The reported activities were also correlated to the accelerometer data-graphs to determine face validity of the measurements at home.

On Monday, the feasibility and comfort of the accelerometers was evaluated and the Pediatric Evaluation of Disability Inventory (PEDI) measuring the performance and capability (self-care, mobility and transfers and functioning) in the activities of daily life of children up to 7.5 years of age, was performed by phone.<sup>19, 24</sup> Since most children with mitochondrial disease are severely limited in their daily activities, we used the PEDI for all age groups.

## Study protocol for healthy controls

The healthy controls were recruited at two regular schools in the surroundings of Nijmegen. Healthy controls were eligible for inclusion when they were healthy and aged between 4 and 18 years. Exclusion criteria: i) confirmed diagnosis of Attention Deficit and Hyperactivity Disorder (ADHD); ii) symptoms of exercise intolerance, fatigue or muscle problems; or iii) the child was under regular surveillance of a paediatrician. Controls were gender- and age matched to a single patient.

Healthy controls were instructed in their home-environment in the same weekend as the age- and gender-matched patient. The attachment and localisation of the accelerometers was the same as the patient protocol. Validation, using the validation protocol and videotaping, was also similar to the patient protocol. Healthy controls were also instructed to wear the accelerometer during waking hours and to keep an activity diary. On Monday, the feasibility and comfort of the accelerometers was evaluated.

## Analyses

### Feasibility

Feasibility was tested using the parent reported complications of wearing the accelerometers and the quantity of the data obtained (% of subjects; the time the device collected data as a percentage of the intended measurement period (Saturday 0:00 – Sunday 23:59) and the time the device collected data as a percentage of the time the sensor was worn). Only patients in whom more than one sensor failed were excluded from the analyses. For the patients in which one sensor failed, only the available data are presented.

### Validity

Validity was assessed by visually correlating the videos with the obtained accelerometer data-graphs (correspondence of orientation and intensity for each sensor) during the validation protocol in each subject. Subsequently, the data from the diaries was correlated to the accelerometer data-graphs (correspondence of the intensity of the movements during the described activities). Only when the video images and described activities clearly did not correlate to the data-graphs, the data were excluded from the analyses. We assessed whether the percentage of dynamic activity and the total leg activity was lower in non-ambulatory children compared to ambulatory children. Finally, the functional abilities, assessed by the GMFM and the PEDI were correlated to the measurements.

## Patients versus controls

We compared patients and their age- and gender matched controls on each of the above mentioned variables. On an individual level, only the percentage of rest was compared between patients and their age and gender matched controls.

## Test-retest reliability

Test-retest reliability was assessed by asking five patients and five age- and gender matched controls to wear the sensors in another season (summer versus spring).

## Covariates

The presence of covariates was assessed by interviewing parents (hours of sports and screen time per week, level of education of parents, hours of sports and body mass index (BMI) of the parents) or by the weather report (Buienradar, Weerplaza). For the analyses assessing the influence of weather conditions, both the results of the test and the retest were used. The following cut-off values were used to indicate 'good weather conditions (e.g. to play outside)': less than 5 mm of rain over the whole weekend; average maximum temperature between 15 and 25 degrees Celsius; and average more than 6 hours of sun per day. Good weather was defined as matching 2 out of 3 criteria; perfect weather was defined as matching all criteria. The influence of these covariates was on the following parameters for activity was assessed: percentage of rest, percentage of dynamic activity, average activity of the upper leg and maximal activity of the upperleg in 30 minutes.

## Statistical analyses

Because of the relatively small number of subjects included in our study, we used non-parametric tests to assess differences and correlations. Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website (<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>). We used a *p*-value of 0.05 for statistical significance. Because of the small numbers and the exploratory character of this study, we did not use Bonferroni adjustment. All analyses were performed using IBM's SPSS statistics software packages, version 20.0.0.1.

## Ethics

This study was approved by the regional Medical Research Ethics Committee (MREC NL50560.091.14). In accordance with the Helsinki agreement, written informed consent was obtained from participant's legal guardian and, where indicated, the participant.

## Results

### Study population

Seventeen patients and 16 healthy age- and gender-matched controls were included in this study from February to May 2015. One healthy control withdrew his consent one day before the measurement would start and no other age- and gender matched control was available for that weekend. The groups were comparable with respect to age, gender, BMI and sports- and highest education of parents, but – as expected – differed significantly with respect to height, weight, time spent at sports and the level of education of the child (Table 1), as well as for the BMI of father. There was a wide variability in the genetic, biochemical, clinical and functional abilities in the children with mitochondrial disease (Supplementary Document 6).

**Table 1: Characteristics of patients and their age- and gender-matched controls.**

	Patients (n = 17)	Healthy controls (n = 15)	
	Median (range)	Median (range)	Difference between groups (p-value)
Gender (male)	8	6	0.69
Age, years	13 (7 - 16)	13 (8 - 17)	0.58
Height, cm	141 (123 - 182)	163 (133 - 188)	0.01
Weight, kg	35 (22 - 75)	52 (30 - 74)	0.01
BMI (kg/m <sup>2</sup> )	17.4 (11.7 - 25.0)	19 (14.2 - 24.4)	0.26
Sports, min/week	0 (0 - 240)	300 (60 - 720)	<0.001
Level of education			
Primary education	2	3	<0.001
Lower general secondary education	1	1	
Higher to a-level general secondary education	1	11	
Special education	13	0	
Screen time, min/day	90 (0 - 300)	150 (60 - 210)	0.91
BMI father (kg/m <sup>2</sup> )	27.0 (22.0 - 39.2)	23.4 (20.8 - 31.2)	0.004
BMI mother (kg/m <sup>2</sup> )	26 (20.5 - 41.8)	24.0 (19.6 - 28.0)	0.15
Completed education father			
Intermediate vocational	10	3	0.05
Higher vocational	6	6	
University	1	5	
Completed education mother			
Intermediate vocational	10	7	0.13

Higher vocational	7	4	
University	0	3	
Sports father, min/week	0 (0 - 240)	150 (0 - 360)	0.06
Sports mother, min/week	60 (0 - 180)	120	0.13
Significant p-values are indicated in bold; BMI = Body Mass Index			

## Feasibility

All participants, including patients with severe mental retardation, tolerated wearing the accelerometers for the duration of the measurement. One patient experienced squeezing of the attachment-top for the chest sensor. The full study protocol was completed by 29 children (88% of total study population): three participants temporarily removed the sensors: two removed the top to ventilate after exercise and because of a party, one did not attach the chest sensor on Sunday because of discomfort of the top and one boy lost his upper leg sensor during outdoor playing. Five sensors failed to record any data and the batteries of one sensors failed during the measurements (18% of all participants; 4% of all measurements). Due to these technical issues, 6% of the total measured time and 8% of the time the sensors were worn was missing in 6 participants (4 patients and 2 healthy controls). One healthy control had to be excluded from the analyses because he lost his upper leg sensor and his upper arm sensor failed to record any data. Most subjects wore their sensors from the moment they awoke to the moment they undressed for bed; three participants removed all sensors after dinner (averagely sensors were worn 94% of the woken time). The time the sensors were taken off because of swimming, showering or bathing was 1.9%. The time the sensors were not worn because of lack of understanding or lack of motivation was 8.7%. The GMFM and the PEDI were successfully executed in all patients. The Tardieu was feasible in all but one patient, who was not able to relax and lie still.

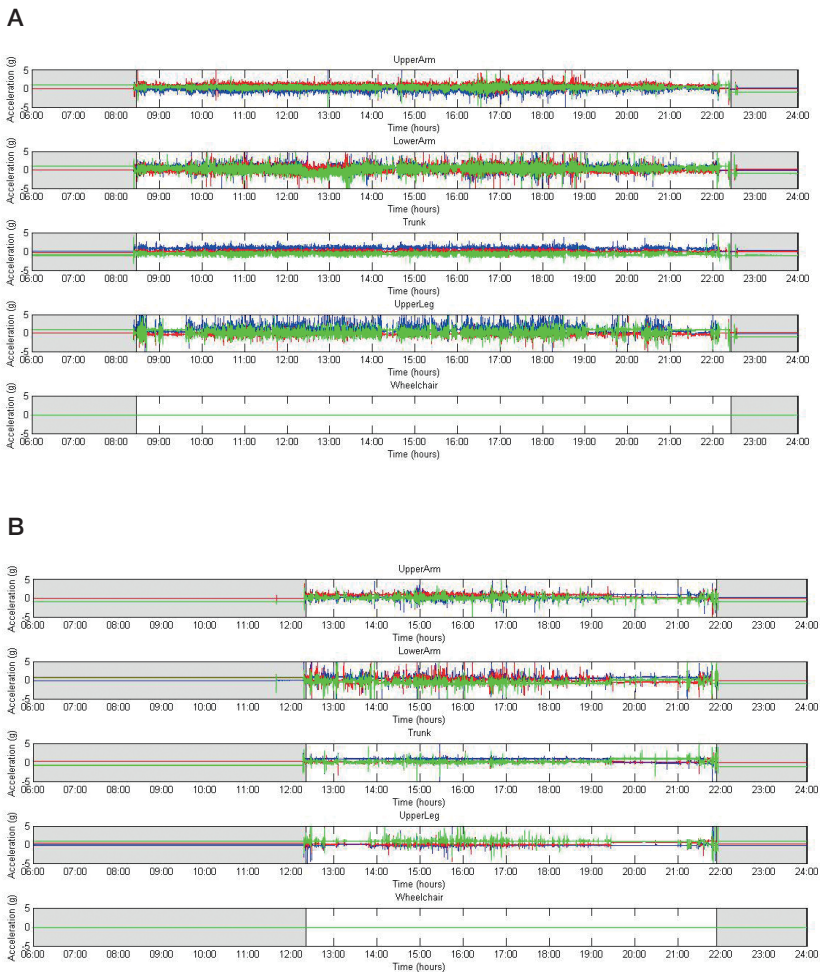
## Validity

For all patients and healthy controls, the movements (orientation, intensity) at the videos corresponded to the acceleration data that was visualized in graphs. In ambulatory versus non-ambulatory patients, dynamic activity (*i.e.* walking) was higher, but the total activity of arms and legs did not reach significance in our small cohort (Table 2;  $p = 0.03$  ( $p$ -level of 0.0042 accepted after Bonferroni correction)). For these analyses, we excluded a boy who was not able to walk but had excellent abilities to move (on his buttock), but not to walk, from these analyses since we could not define in which group he belonged. We found a (very) strong and significant correlation between the motor abilities as measured with the GMFM and the resting percentage ( $\rho = -0.82$ ), the largest amount of activity of the leg during half an hour ( $\rho = 0.65$ ) and the peak-activity of the lower arm ( $\rho = 0.67$ ; all  $p < 0.0001$ ). The score on the mobility domain of the PEDI

(functional abilities) also correlated very strongly with the resting percentage ( $\rho = -0.87$ ) and strongly with the largest amount of activity of the leg during half an hour ( $\rho = 0.70$ ) and the peak-activity of the leg ( $\rho = 0.60$ ; all  $p < 0.0001$ ; Supplementary Document 7).

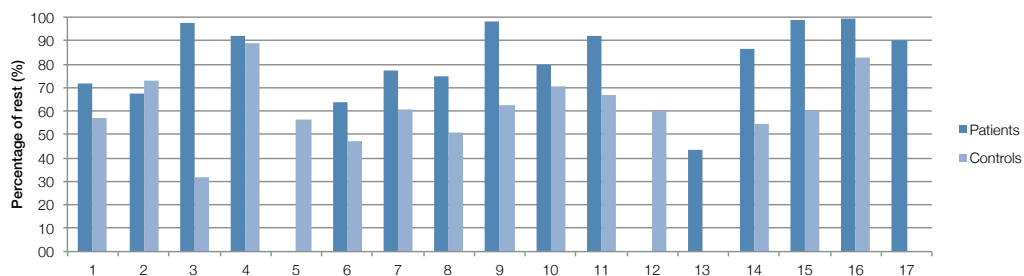
### Patients versus controls

The patients with mitochondrial disorders in our cohort had lower maximal intensity and were resting more compared to their age- and gender-matched controls (Table 2; Figure 2).



**Figure 2.**

**Raw data for a matched couple.** The couple (couple 3) with the most striking difference was selected, the healthy control is in Figure 2A and Figure 2B represents the patient.



**Figure 3.**

**Percentage of rest during the weekend for patients versus matched controls.** The healthy girl of couple 2 had to study for her final exams during the weekend. The patient in couple 3 is a boy with behavioural problems and hyperactivity who removed the sensors during his afternoon nap (percentage of rest is expressed as a percentage of the hours the accelerometers were worn). The dark bar represent the patients and the light bars represent their matched controls.

We saw no difference in arm activity between patients and controls. When comparing individual results, all but one patient had higher percentages of rest during the weekend, except for one girl matched to a healthy girl who had to study for her final exams during the whole weekend (couple 2 in Figure 3).

### Test-retest reliability

Five patients and five healthy controls were included in the retest from June to July 2015. Two patients were matched to another age- and gender matched control because of the healthy control was not available in the same weekend as the patient. One patient refused to participate because of the weather conditions (exceptionally warm for the Netherlands: 35°C with high humidity); her healthy control was therefore also excluded from the analyses. We experienced no technical failures during the retest. The retest was averagely executed 2.4 months later (spread 1.7 – 3.5 months), always in a different season (spring versus summer). The children selected for the retest were comparable with respect to age, BMI, screen time, time spent at sports activities, as well as PEDI and GMFM score ( $p = 0.05 - 0.8$ ). The weather was not comparable with respect to temperature and hours of sun ( $p = 0.012$  and  $p = 0.036$ , respectively; higher in summer), though the amount of rain was comparable ( $p = 0.13$ ). The time the sensors were worn at the retest were not significantly different from the first test ( $p = 0.48$ ). We found that the activity (activity level, maximal intensity and largest AUC during 30 minutes) in summer were not significantly different from the activity in spring ( $p = 0.21 - 1.0$ ; Table 3).



Table 2.

**Average activity over the weekend for children with mitochondrial disease compared to age- and gender-matched controls.** We excluded a boy who was not able to walk but had excellent abilities to move (on his buttock), but not to walk, from the ambulatory/non-ambulatory analyses since we could not define in which group he belonged.

	Patients (n = 17)		Healthy controls (n = 15)		P-value for difference between patients and healthy controls	Ambulatory patients (n = 9)		Non ambulatory patients (n = 7)		P-value for difference between ambulatory and non-ambulatory patients	Myopathy (n = 7)		Encephalo-myopathy (n = 10)		P-value for difference between patients with myopathy and patients with encephalo-myopathy
	Median	Range	Median	Range		Median	Range	Median	Range		Median	Range	Median	Range	
Wear time (hours per day)	10	(8 - 13)	12	(6 - 15)	<b>0.004</b>	12	(9 - 14)	11	(9 - 12)	0.54	11	(8 - 13)	10	(8 - 12)	0.28
Average % rest	86	(44 - 99)	60	(2 - 89)	<b>&lt;0.001</b>	73	(44 - 90)	98	(86 - 99)	<b>0.006</b>	77	(63 - 99)	89	(44 - 98)	0.6
Average % standing	4	(0 - 23)	18	(7 - 48)	<b>0.004</b>	16	(4 - 23)	1	(0 - 3)	<b>&lt;0.001</b>	10	(0 - 23)	3	(0 - 22)	0.8
Average % dynamic activity	13	(1 - 50)	18	(6 - 33)	<b>0.02</b>	14	(3 - 34)	5	(1 - 50)	0.09	13	(1 - 50)	9	(6 - 35)	0.4
Average counts upper arm (1,000 counts/hour)	9.4	(3.1 - 29.6)	18.0	(7.0 - 27.2)	<b>0.008</b>	12.7	(8.2 - 29.6)	5.5	(3.1 - 17.4)	<b>0.03</b>	9.1	(3.1 - 16.9)	10.0	(5.1 - 29.6)	0.7
Average counts upper arm (1,000 counts/hour)	15.7	(11.3 - 36.2)	19.0	(10.5 - 30.2)	0.3	18.9	(14.2 - 32.2)	14.3	(11.3 - 36.2)	0.10	16.5	(11.3 - 25.2)	15.5	(13.2 - 36.2)	0.6
Average counts lower arm (1,000 counts/hour)	21.5	(16.1 - 41.8)	27.1	(15.3 - 37.1)	0.3	23.3	(18.9 - 40.3)	20.0	(16.1 - 41.8)	0.10	21.5	(16.1 - 31.8)	21.2	(17.8 - 41.8)	0.9
Average counts wheelchair (1,000 counts/hour)	1.3	(0.7 - 1.4)	NA	NA	NA	NA	0.8	1.3	(0.7 - 1.5)	0.60	NA	1.4	1.2	(0.7 - 1.5)	0.9

Maximal intensity upper leg (1,000 counts/min)	0.97 (0.5 - 2.6)	1.9 (0.9 - 2.6)	<b>&lt;0.001</b>	1.3 (0.7 - 2.6)	0.9 (0.5 - 1.0)	<b>0.03</b>	0.9 (0.5 - 1.7)	1.0 (0.9 - 2.6)	0.6
Maximal intensity upper arm (1,000 counts/min)	0.97 (0.8 - 1.9)	1.6 (0.9 - 2.1)	<b>&lt;0.001</b>	1.1 (0.8 - 1.9)	0.9 (0.8 - 1.2)	0.06	1.0 (0.9 - 1.4)	1.0 (0.8 - 1.9)	0.8
Maximal intensity lower arm (1,000 counts/min)	1.2 (0.9 - 2.2)	1.9 (1.0 - 2.5)	<b>&lt;0.001</b>	1.4 (1.0 - 2.2)	1.1 (0.9 - 1.4)	0.09	1.2 (1.0 - 1.8)	1.2 (0.9 - 2.2)	0.9
Largest AUC during ½ hour upper leg (1,000 counts)	10.4 (6.4 - 37.3)	26.8 (11.4 - 66.2)	<b>&lt;0.001</b>	20.2 (8.9 - 37.2)	9.9 (6.4 - 17.3)	<b>0.03</b>	9.5 (6.4 - 37.3)	13.9 (7.4 - 30.7)	0.4
Largest AUC during ½ hour upper arm (1,000 counts)	15.8 (1.1 - 27.3)	25.3 (11.7 - 368)	<b>0.003</b>	17.9 (14.0 - 27.3)	14.6 (11.1 - 25.1)	0.09	15.5 (11.1 - 27.3)	17.0 (11.9 - 25.1)	0.5
Largest AUC during ½ hour lower arm (1,000 counts)	21.2 (16.0 - 34.8)	30.8 (16.7 - 43.4)	<b>0.004</b>	25.6 (18.5 - 34.8)	18.2 (16.0 - 29.3)	<b>0.04</b>	20.4 (16.0 - 34.8)	22.9 (16.7 - 29.9)	0.5

Significant *p*-values are indicated in bold; AUC = Area under the curve; min = minutes

**Table 3.**

**Test-retest in four patients and four controls.** One patient refused to wear the accelerometers because of warm weather (35 °C).

	First weekend	Second weekend	P-value for difference between first and second weekend
	Median (range)	Median (range)	
Wear time (hours per day)	11 (9 - 14)	10 (6 - 15)	0.48
Sun (hours)	8 (2 - 10)	13 (7 - 18)	<b>0.002</b>
Temperature (°C)	34 (13 - 25)	33 (30 - 37)	<b>0.001</b>
Amount of rain (mm)	2 (0 - 9)	0 (0 - 4)	0.9
Average % rest	75 (32 - 98)	70 (55 - 98)	0.5
Average % standing	5 (0 - 48)	14 (0 - 32)	0.8
Average % dynamic activity	14 (2 - 33)	13 (2 - 21)	0.6
Average counts upper leg (1.000 counts/hour)	12.4 (5.1 - 27.2)	12.6 (4.8 - 21.6)	0.6
Average counts upper arm (1.000 counts/hour)	16.6 (13.8 - 25.0)	16.3 (11.5 - 22.3)	0.3
Average counts lower arm (1.000 counts/hour)	21.9 (17.8 - 34.5)	22.5 (11.9 - 30.8)	0.7
Maximal intensity upper leg (1.000 counts/min)	1.3 (0.9 - 2.5)	1.3 (0.7 - 3.3)	0.7
Maximal intensity upper arm (1.000 counts/min)	1.3 (0.8 - 1.9)	1.3 (0.7 - 2.1)	0.3
Maximal intensity lower arm (1.000 counts/min)	1.6 (0.9 - 2.5)	1.3 (1.0 - 2.5)	0.7
Largest AUC during ½ hour upper leg (1.000 counts)	17.1 (9.9 - 45.8)	18.0 (7.4 - 23.2)	0.2
Largest AUC during ½ hour upper arm (1.000 counts)	17.5 (11.9 - 30.7)	19.5 (11.1 - 24.2)	1.0
Largest AUC during ½ hour lower arm (1.000 counts)	21.8 (16.7 - 36.2)	21.6 (16.0 - 33.4)	0.8

Significant *p*-values are indicated in bold; AUC = Area under the curve; min = minutes

Except for the temperature ( $p = 0.02$ ), the weather between the Saturdays and the Sundays was comparable (Table 4). Comparing Saturday to Sunday, we only found significantly lower arm activity in patients on Sunday compared to Saturday (21.1 vs 21.9; Table 4). Sensors were worn for more hours on Saturdays compared to Sundays ( $p = 0.002$  for both groups).

## Covariates

Good or perfect weather was not associated with increased activity ( $p = 0.25 - 0.74$  and  $p = 0.37 - 0.71$ , respectively). Activity was also not different between weekends with more than 5 mm of rain, compared to those with less than 5 mm of rain ( $p = 0.36 - 0.81$ ).

Maximal intensity upper arm (1,000 counts/min)	1.0 (0.7 - 1.9)	1.0 (0.8 - 1.5)	0.7	1.4 (0.8 - 2.1)	1.2 (0.8 - 1.9)	0.6
Maximal intensity lower arm (1,000 counts/min)	1.1 (0.9 - 2.2)	1.2 (0.9 - 1.8)	0.6	1.8 (1.0 - 2.8)	1.5 (1.0 - 2.5)	0.3
Largest AUC during ½ hour upper leg (1,000 counts)	10.4 (6.4 - 37.3)	9.9 (1.3 - 18.7)	0.5	24.0 (11.4 - 66.2)	19.0 (6.0 - 63.2)	0.7
Largest AUC during ½ hour upper arm (1,000 counts)	15.3 (10.5 - 27.3)	15.2 (11.1 - 26.4)	0.5	23.4 (11.7 - 25.8)	18.3 (8.0 - 36.8)	0.6
Largest AUC during ½ hour lower arm (1,000 counts)	20.4 (12.7 - 34.8)	19.4 (15.8 - 32.8)	0.7	37.1 (16.5 - 40.6)	24.8 (11.7 - 43.4)	0.2

A level of  $p = 0.0042$  was used for significance; significant  $p$ -values are indicated in bold; AUC = Area under the curve; min = minutes

**Table 4.**  
**Activity on Saturday versus Sunday for patients and healthy controls.**

	Patients (n = 17) Saturday		Sunday	p-value for difference between Saturday and Sunday in patients	Healthy controls (n = 17) Saturday		Sunday	p-value for difference between Saturday and Sunday in healthy controls
	Median (range)	Median (range)	Median (range)		Median (range)			
Worn hours	11 (8 - 16)	10 (4 - 13)	0.002	13 (10 - 18)	11 (7 - 15)	0.002		
Sun (hours)	2.5 (0.8 - 10.0)	5.6 (0.0 - 13.7)	0.1	2.8 (0.0 - 7.6)	7.6 (0.0 - 13.4)	0.1		
Temperature (°C)	15.5 (7.4 - 17.0)	16.5 (6.4 - 19.1)	0.02	15.4 (5.5 - 17.7)	17.1 (5.5 - 17.7)	0.02		
Amount of rain (mm)	0.8 (0 - 5.1)	0 (0.0 - 20.6)	0.6	1.6 (0.0 - 6.3)	0 (0.0 - 15.3)	0.4		
Average % rest	83 (51.7 - 99.2)	89 (35.4 - 99.3)	0.5	65 (35 - 84)	54 (28 - 94)	0.7		
Average % standing	1 (0.0 - 22.3)	5 (0.0 - 25.5)	0.7	17 (7 - 46)	18 (3 - 49)	0.4		
Average % dynamic activity	6 (0.8 - 30.6)	6 (0.7 - 39.0)	0.7	17 (6 - 27)	22 (3 - 51)	0.2		
Average counts upper leg (1,000 counts/hour)	9.7 (2.9 - 29.3)	9.2 (3.4 - 30.0)	0.2	15.2 (7.8 - 23.6)	18.2 (5.6 - 40.2)	0.4		
Average counts upper arm (1,000 counts/hour)	16.6 (10.3 - 36.5)	16.2 (12.0 - 36.0)	0.4	17.8 (11.2 - 25.4)	20.3 (9.3 - 35.0)	0.2		
Average counts lower arm (1,000 counts/hour)	21.9 (14.9 - 41.3)	21.1 (14.0 - 42.3)	0.03	26.0 (14.1 - 32.6)	28.0 (13.6 - 41.5)	0.2		
Average counts wheelchair (1,000 counts/hour)	2.6 (0.5 - 3.2)	2.2 (1.2 - 3.2)	0.9	NA	NA	NA		
Maximal intensity upper leg (1,000 counts/min)	1.0 (0.5 - 2.6)	0.9 (4.3 - 29.4)	0.4	1.0 (0.8 - 2.5)	1.4 (0.8 - 2.4)	0.9		

## Discussion

In this study, we showed that measuring physical activity in a home-situation with 3D accelerometers was feasible in all seventeen children with mitochondrial disorders, although we experienced technical difficulties with the hardware in 18% of the subjects (4% of the measurements). We found good face validity by visually correlating the validation videos and the diaries to the data-graphs. Moreover, non-ambulatory patients had lower dynamic activity (*i.e.* walking) compared to ambulatory patients, although the leg activity level only showed a trend towards lower activity levels in non-ambulatory children. Generally, patients with mitochondrial disorders had lower maximal intensity and were resting more, compared to their age- and gender-matched peers. A retest was performed in eight patients in a weekend with significantly different weather and showed good test-retest reliability.

A previous study in six children with mitochondrial disease showed a lower activity level compared to healthy controls and less time spent in moderate to vigorous activities.<sup>25</sup> Our cohort included both schoolchildren and adolescents, of which the latter are known to be much less active compared to schoolchildren.<sup>26</sup> Unexpectedly, arm activity levels were comparable between patients and healthy controls. Our data could not confirm that this was due to compensatory use of arms in non-ambulatory children or high levels of unpurposeful arm activity in children with movement disorders. It could also be suggested that arm activity is normal in children with mitochondrial disorders, or that our method was not sensitive enough to detect any differences. The resting percentage in the study by Martens *et al.* was lower compared to what we found, even in ambulatory patients only. This could be due to selection bias: Martens *et al.* studied ambulatory children without severe cognitive impairment and none of them had a genetically confirmed mitochondrial disease.

Objectively measuring daily physical activity in a patient's home environment was feasible using a 3D accelerometer in five young boys (4 - 6 years) with Duchenne Muscular Dystrophy (DMD).<sup>13</sup> In this study, the accelerometer was able to reliably measure body posture, walking parameters and the distribution of these activities over the day. These ambulatory boys, although significantly younger than our study subjects, had comparable levels of dynamic activity, but much less time spent resting, compared to our ambulatory patients.<sup>13</sup> In patients with cerebral palsy (CP) with different levels of disability, no adverse effects or discomfort were reported when wearing one accelerometer.<sup>27</sup> The average time spent resting for ambulatory children in our cohort was comparable to the resting time of children with CP GMFCS level I and II. Non-ambulatory patients were comparable to GMFCS level IV patients. The percentage of resting in ambulatory normally-weighted children with Downs, Williams and Prader-Willi syndrome was comparable to our ambulatory patients.<sup>28</sup>

During a retest in eight during weekends with significantly different temperature and hours of sun, none of the parameters (activity level, maximal intensity and largest AUC during 30 minutes) was significantly different. In healthy adults, test-retest reliability of accelerometry was reported to be very good to excellent and there were no differences in activity measured for 7-day periods, 1 to 4 weeks apart.<sup>29</sup> The activity pattern of healthy children is more stable over the day during weekends.<sup>26, 30</sup> Activity is known to be significantly lower in winter,<sup>26</sup> mainly during rainfall and cold and windy weather.<sup>31</sup> In our small and heterogeneous sample, we could not detect these differences. During maturation, day-to-day variability in activity and the seasonal influence on activity decrease.<sup>32</sup> We found no correlation between age and activity nor higher activity in children under 12 years (both in the patient group and the healthy control group).

Whereas most accelerometers are attached to the hip and only report the level of activity, we used a set of accelerometers at the chest, the lower arm and upper arm, upper leg and if present the wheelchair, to assess both the orientation and the intensity of movements of these body parts. In the future, this measurement protocol will allow us to customize our analyses to increase the sensitivity of the measurements. These analyses might for example include characterization of gait pattern, but also the parallel measurement of physiological parameters could be used to indicate the effort of movements, which might increase the sensitivity of the measurement, but also requires more sophisticated methods for data-analysis. Alternatives include the use of a simple commercially available activity monitor, that provides insight in the overall activity of a person. Although such a monitor may be more user-friendly in use, it most probably doesn't allow specific tailor-made analyses. Another alternative would be to make the current method more user-friendly by reducing the number of sensors and/or improving the convenience of using the script.

Strengths of this study include the large and clinically heterogeneous group of patients with genetically confirmed mitochondrial disease in this study, which allows for conclusions about the feasibility of the accelerometers in both patients with high and with limited abilities. Secondly, a healthy age- and gender-matched population was included at the same weekend as the patient was measured to minimize the influence of weather conditions. Finally, the influence of season and weather to the activity of both patients and their peers was assessed in a formal but small test-retest reliability study.

In contrast to most previous studies, we also included non-ambulatory patients in our study. Measuring daily activity in these patients is challenging, because accelerations elicited by the wheelchair are also measured by the other sensors. We found that the wheelchair accounted for about one fourth of the leg activity level in non-ambulatory children. Other methodological issues include the lack of measurement of stable, sustained body positions since accelerometers measure acceleration and that

the current data analysis is not able to differentiate between meaningful movements and aberrant movements, such as myoclonus or ataxia. In our patient cohort, none of the arm-parameters was significantly higher in patients with ataxia. Other weaknesses include the technical challenges that still remain, including the failure of the batteries. Finally, we only measured physical activity during two days, whereas measuring for longer periods decreased variability substantially in patients with CP.<sup>33</sup>

In the near future, we will optimize the technical issues (*i.e.* the MOX-accelerometer battery) and analysis methods. The latter will include the validation of an analysis to correct for wheelchair movements and more detailed methods to quantify activity. Finally, we will define a more detailed and clinically relevant end point in close dialogue with patients and their families. For example, since patients frequently report gait disturbances and balance problems<sup>2</sup> especially when they are tired, we might focus on analysing gait disturbances.

In conclusion, measuring daily activity with 3D accelerometers was practically feasible in all seventeen patients with mitochondrial disease, although we were confronted with hardware failures in 18% of the cases. Sensors were worn approximately 94% of the woken time. Validity was good and test-retest reliability looks promising. When using the accelerometer in a future clinical trial, we suggest including a longer measurement period and standardization of family activities during the measurements, to provide representative activity patterns and reduce random variability. Although compliance was good for this study, some adaptations can be made to make wearing an accelerometer more acceptable.<sup>34</sup> In the near future, we will optimize our measurement protocol by defining more sensitive and clinically relevant outcome parameters in close dialogue with patient and parents.

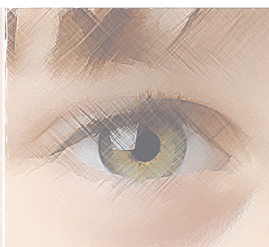
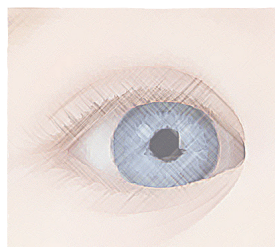
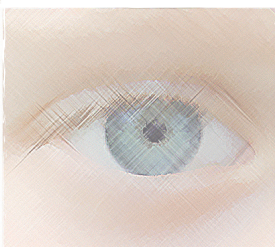
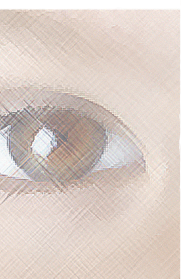
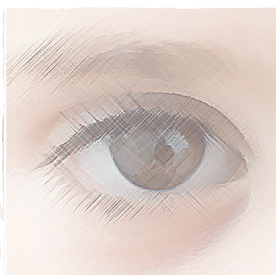
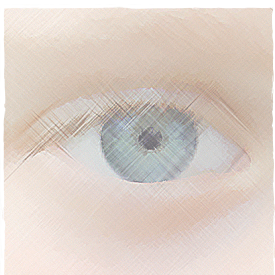
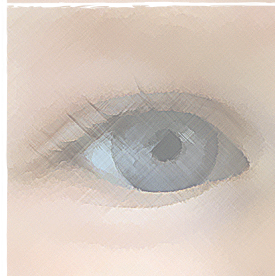
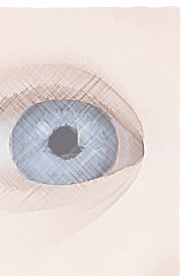
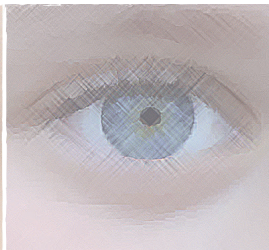
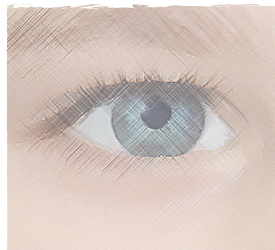
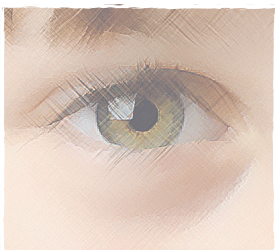


## References

1. Pfeffer G, Horvath R, Klopstock T, *et al.* New treatments for mitochondrial disease-no time to drop our standards. *Nature reviews Neurology* 2013;9:474-481.
2. Koene S, Wortmann SB, de Vries MC, *et al.* Developing outcome measures for pediatric mitochondrial disorders: which complaints and limitations are most burdensome to patients and their parents? *Mitochondrion* 2013;13:15-24.
3. Gordijn SM, Cremers EM, Kaspers GJ, Gemke RJ. Fatigue in children: reliability and validity of the Dutch PedsQL(TM) Multidimensional Fatigue Scale. *Qual Life Res* 2011;20:1103-1108.
4. Swanink CM, Vercoulen JH, Bleijenberg G, Fennis JF, Galama JM, van der Meer JW. Chronic fatigue syndrome: a clinical and laboratory study with a well matched control group. *Journal of internal medicine* 1995;237:499-506.
5. LeMura LM, von Duvillard SP, Cohen SL, *et al.* Treadmill and cycle ergometry testing in 5- to 6-year-old children. *European journal of applied physiology* 2001;85:472-478.
6. Beenakker EA, Maurits NM, Fock JM, Brouwer OF, van der Hoeven JH. Functional ability and muscle force in healthy children and ambulant Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2005;9:387-393.
7. Parreira SL, Resende MB, Zanoteli E, Carvalho MS, Marie SK, Reed UC. Comparison of motor strength and function in patients with Duchenne muscular dystrophy with or without steroid therapy. *Arq Neuropsiquiatr* 2010;68:683-688.
8. Abel MF, Damiano DL, Blanco JS, *et al.* Relationships among musculoskeletal impairments and functional health status in ambulatory cerebral palsy. *J Pediatr Orthop* 2003;23:535-541.
9. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports* 1985;100:126-131.
10. Bjornson KF. Physical activity monitoring in children and youths. *Pediatr Phys Ther* 2005;17:37-45.
11. Capio CM, Sit CH, Abernethy B. Physical activity measurement using MTI (actigraph) among children with cerebral palsy. *Arch Phys Med Rehabil* 2010;91:1283-1290.
12. Clanchy KM, Tweedy SM, Boyd RN, Trost SG. Validity of accelerometry in ambulatory children and adolescents with cerebral palsy. *European journal of applied physiology* 2011;111:2951-2959.
13. Jeannet PY, Aminian K, Bloetzer C, Najafi B, Paraschiv-Ionescu A. Continuous monitoring and quantification of multiple parameters of daily physical activity in ambulatory Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2011;15:40-47.
14. McDonald CM, Widman L, Abresch RT, Walsh SA, Walsh DD. Utility of a step activity monitor for the measurement of daily ambulatory activity in children. *Arch Phys Med Rehabil* 2005;86:793-801.
15. Clanchy KM, Tweedy SM, Boyd R. Measurement of habitual physical activity performance in adolescents with cerebral palsy: a systematic review. *Developmental medicine and child neurology* 2011;53:499-505.
16. Koene S, Jansen M, Verhaak CM, De Vroeh RL, De Groot IJ, Smeitink JA. Towards the harmonization of outcome measures in children with mitochondrial disorders. *Developmental medicine and child neurology* 2013.
17. Meijer K, Annegarn J, Lima Passos V, *et al.* Characteristics of daily arm activities in patients with COPD. *Eur Respir J* 2014;43:1631-1641.
18. Nordmark E, Jarnlo GB, Hagglund G. Comparison of the Gross Motor Function Measure and Paediatric Evaluation of Disability Inventory in assessing motor function in children undergoing selective dorsal rhizotomy. *Developmental medicine and child neurology* 2000;42:245-252.

19. Vos-Vromans DC, Ketelaar M, Gorter JW. Responsiveness of evaluative measures for children with cerebral palsy: the Gross Motor Function Measure and the Pediatric Evaluation of Disability Inventory. *Disabil Rehabil* 2005;27:1245-1252.
20. Russell D, Rosenbaum, PL., Avery, LM., Lane, M. *Gross Motor Function Measure (GMGM-66 and GMFM-88) User's Manual*. London, United Kingdom: Mac Keith Press 2002.
21. Brunton LK, Bartlett DJ. Validity and reliability of two abbreviated versions of the Gross Motor Function Measure. *Phys Ther* 2011;91:577-588.
22. Gracies JM, Burke K, Clegg NJ, *et al*. Reliability of the Tardieu Scale for assessing spasticity in children with cerebral palsy. *Arch Phys Med Rehabil* 2010;91:421-428.
23. Scholtes VA, Becher JG, Beelen A, Lankhorst GJ. Clinical assessment of spasticity in children with cerebral palsy: a critical review of available instruments. *Developmental medicine and child neurology* 2006;48:64-73.
24. Custers JW, Wassenberg-Severijnen JE, Van der Net J, Vermeer A, Hart HT, Helders PJ. Dutch adaptation and content validity of the 'Pediatric Evaluation Of Disability Inventory (PEDI)'. *Disabil Rehabil* 2002;24:250-258.
25. Martens AM, Gorter H, Wassink RG, Rietman H. Physical activity of children with a mitochondrial disease compared to children who are healthy. *Pediatr Phys Ther* 2014;26:19-26.
26. Riddoch CJ, Mattocks C, Deere K, *et al*. Objective measurement of levels and patterns of physical activity. *Archives of disease in childhood* 2007;92:963-969.
27. Gorter JW, Noorduyn SG, Obeid J, Timmons BW. Accelerometry: a feasible method to quantify physical activity in ambulatory and nonambulatory adolescents with cerebral palsy. *International journal of pediatrics* 2012;2012:329284.
28. Nordstrom M, Hansen BH, Paus B, Kolset SO. Accelerometer-determined physical activity and walking capacity in persons with Down syndrome, Williams syndrome and Prader-Willi syndrome. *Res Dev Disabil* 2013;34:4395-4403.
29. Sirard JR, Forsyth A, Oakes JM, Schmitz KH. Accelerometer test-retest reliability by data processing algorithms: results from the Twin Cities Walking Study. *Journal of physical activity & health* 2011;8:668-674.
30. Brooke HL, Corder K, Atkin AJ, van Sluijs EM. A systematic literature review with meta-analyses of within- and between-day differences in objectively measured physical activity in school-aged children. *Sports medicine (Auckland, NZ)* 2014;44:1427-1438.
31. Chan CB, Ryan DA. Assessing the effects of weather conditions on physical activity participation using objective measures. *International journal of environmental research and public health* 2009;6:2639-2654.
32. Rich C, Griffiths LJ, Dezateux C. Seasonal variation in accelerometer-determined sedentary behaviour and physical activity in children: a review. *The international journal of behavioral nutrition and physical activity* 2012;9:49.
33. Mitchell LE, Ziviani J, Boyd RN. Variability in measuring physical activity in children with cerebral palsy. *Medicine and science in sports and exercise* 2015;47:194-200.
34. Kirby J, Tibbins C, Callens C, *et al*. Young People's Views on Accelerometer Use in Physical Activity Research: Findings from a User Involvement Investigation. *ISRN obesity* 2012;2012:948504.





# Chapter 7

## The International Paediatric Mitochondrial Disease Scale



## The International Paediatric Mitochondrial Disease Scale

**Saskia Koene**<sup>1</sup>, Jan C.M. Hendriks<sup>2</sup>, Ilse Dirks<sup>1</sup>, Lonneke de Boer<sup>1</sup>, Maaïke C. de Vries<sup>1</sup>, Mirian C.H. Janssen<sup>1</sup>, Izelle Smuts<sup>3</sup>, Cheuk-Wing Fung<sup>4</sup>, Virginia C.N. Wong<sup>4</sup>, I. (René) F.M. de Co<sup>5</sup>, Katharina Vill<sup>6</sup>, Claudia Stendel<sup>6</sup>, Thomas Klopstock<sup>7,8</sup>, Marni J. Falk<sup>9</sup>, Elizabeth M. McCormick<sup>10</sup>, Robert McFarland<sup>11</sup>, Imelda J.M. de Groot<sup>1,12</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders at the Department of Paediatrics, Radboudumc

<sup>2</sup>Department of Health Evidence, Radboudumc

<sup>3</sup>Steve Biko Academic Hospital, Ludwig-Maximilians-of Pretoria, South Africa

<sup>4</sup>Department of Paediatrics & Adolescent Medicine, The University, Munich, Germany of Hong Kong

<sup>5</sup>Department of Neurology, ErasmusMC, Rotterdam, the Netherlands

<sup>6</sup>Department of Neurology, Friedrich-Baur-Institute, Ludwig-Maximilians-University, Munich, Germany

<sup>7</sup>Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

<sup>8</sup>German Center for Neurodegenerative Diseases (DZNE), Munich, Germany

<sup>9</sup>Division of Human Genetics, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

<sup>10</sup>Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA

<sup>11</sup>Wellcome Trust Centre for Mitochondrial Research Newcastle

<sup>12</sup>Department of Rehabilitation, Radboudumc

Submitted

There is an urgent need for reliable and universally applicable outcome measures for children with mitochondrial diseases. In this study, we aimed to adapt the currently available Newcastle Paediatric Mitochondrial Disease Scale to the International Paediatric Mitochondrial Disease Scale (IPMDS) during a Delphi-based process with input from international collaborators, patients and caretakers as well as a pilot reliability study in eight patients. Subsequently, we aimed to test the feasibility, construct validity and reliability of the IPMDS in a multi-centre study.

A clinically, biochemically and genetically heterogeneous group of seventeen patients (age 1.6 -16 years) from five different expert centres from four different continents were included in this study. The feasibility of the IPMDS was good, as indicated by a low number of missing items (4%) and the positive evaluation of patients, parents and users. Principal component analysis of our small sample identified three factors, which explained 57.9% of the variance. Good construct validity was found using hypothesis testing. The overall inter-rater reliability was good (median ICC<sub>agreement</sub> 0.85; range 0.23 – 0.99).

In conclusion, we suggest using the IPMDS for the assessment of natural history in children with mitochondrial diseases. These data should be used to further explore the construct validity of the IPMDS and to set age limits. In parallel, responsiveness and the minimal clinically important difference should be studied to facilitate sample size calculations in future clinical trials.



## Introduction

Mitochondrial diseases are the most prevalent inherited metabolic diseases, with an incidence of approximately 1 in 5,000 live births.<sup>1</sup> Since mitochondria are present in almost all cells, symptoms can arise theoretically from every organ. The most commonly affected organs and tissues include the brain, eye, heart and skeletal muscle.<sup>2</sup> There is enormous variability in the pattern of affected organs and the degree of disability experienced. Whereas some children with mitochondrial disease thrive in mainstream school and live well into adult life, others follow a more rapidly progressive course and die in the neonatal period or function at a low level, barely interacting with their environment.

Currently, there is no cure for mitochondrial diseases, but there are some promising results of pharmacological interventions in cells and animals and the prospects for randomised clinical trials of novel and re-purposed pharmaceuticals are increasing.<sup>2-8</sup> Outcome measures that are valid, reliable, sensitive and clinically relevant are critical to the success of such trials, but the heterogeneity and multi-systemic nature of mitochondrial diseases pose significant challenges in choosing an appropriate, universally applicable outcome measure.<sup>9</sup> To be able to measure disease severity and disease progression within the full range of the phenotypic spectrum, a combination of objective, subjective, functional and biochemical end points will probably be indicated.

A mitochondrial disease specific follow up tool for children already exists, namely the Newcastle Paediatric Mitochondrial Disease Scale (NPMDS).<sup>10</sup> This scale was originally designed to be a concise and pragmatic clinical tool to monitor the biophysical markers of disease progression. Although the NPMDS fits this purpose from a natural disease course perspective, it was not designed as an end point instrument and probably lacks a sufficient level of detail required for this purpose in clinical trials. Moreover, the scale was not developed to measure the clinically relevant concept of functional disability.<sup>10</sup>

In this study, we aim to adapt the NPMDS to a more clinically relevant and more detailed scoring system for future clinical trials in paediatric mitochondrial disease. This International Paediatric Mitochondrial Disease Scale (IPMDS) should cover more of the symptoms indicated by patients and parents as ‘burdensome’, such as tiredness and lack of energy, behavioural problems and depression.<sup>11</sup> Besides, we aimed to include a functional domain to quantify changes in the motor abilities of the child, since clinically relevant changes in motor function are not always equally reflected by changes in muscle power or tone, and vice versa.<sup>12-14</sup> After a Delphi-based development process, we aimed to test the construct validity and reliability (inter-rater, intra-rater and test-retest) by field testing in several international expert centres.

## Methods

The IPMDS was developed during a Delphi-based process, consulting patients, parents and mitochondrial disease experts. After a first pilot reliability test, the scale was further optimised for subsequent testing in five expert centres. At each centre, two to four randomly selected patients were assessed by three to four physicians to evaluate inter-rater reliability. The construct validity was tested using factor analysis and by hypothesis testing. In a subset of patients within these studies, test-retest reliability and intra-rater reliability was tested.

This study was conducted in all countries after approval from the regional Medical Research Ethics Committee (MREC NL.44833.091.13). In accordance with the Helsinki agreement, written informed consent was obtained from each participant or his/her legal guardian(s).

### Development of the scale

The first version of the scoring list was composed and reviewed by a team of six physicians from Nijmegen, which included three paediatricians, one paediatric resident, one paediatric rehabilitation specialist and one internal medicine physician, all seeing patients with mitochondrial diseases on a regular basis. To get a general impression of the functioning of the patient, both subjective and objective items were included. For the subjective part, we chose to interview parents and ask for their opinion on the child's health in the past 4 weeks, similar to the NPMDS. Items in the subjective domain were based on a previous study on which sign and symptoms parents would most like to change.<sup>11</sup> Besides, some items were partly adopted from existing scales, such as the NPMDS,<sup>10</sup> the Newcastle Mitochondrial Disease Adult Scale (NMDAS)<sup>15</sup> and the Paediatric Evaluation of Disability Inventory (PEDI).<sup>16</sup> Items within the objective physical examination were dictated by the relevant general paediatric and neurological examination, including items scored with previously validated scores.<sup>17</sup> Functional tests were inspired both by the prevalence and severity of disabilities within the International Classification of Functioning in Children and Youth (ICF-CY)<sup>11</sup> and the domains of the Motor Function Measure (MFM).<sup>18</sup> Since children under 6 years are not expected to be able to hop, this item (3.10), as well as the running item (3.9) and the rotation of a pen with a single hand (3.13), were left out of the questionnaire for children under 6 years. A manual was constructed to facilitate similar conduction of the items by all researchers.

All items were reviewed critically for completeness, relevance, uniformity, acceptability for patients and practicability by the six physicians from Nijmegen. Missing items were included. The adapted, second version of the scoring list was applied to two children with a mitochondrial disease, one child with a mild phenotype and one child with a severe condition to test for acceptability and practical difficulties. Practical adaptations

were made afterwards and evaluated again by all physicians. To facilitate exclusion of items that were difficult to assess (e.g. headache in a severely intellectually disabled child), it was made possible to adapt the maximum score if an item was impossible to indicate.

The pilot scoring list was sent to all centres participating in this study for revision and adaptations were made on the consensus.

### Pilot study

For the reliability and validity study of the pilot scoring list, eight patients were invited to the Nijmegen Centre for Mitochondrial Disorders outpatient clinic. A heterogeneous group of patients was selected, with the focus on patients in which previously difficulties in performing the NPMDS were experienced. Over the course of one day, patients were scored by four physicians, with a break of 15 minutes between all consultations and a one-hour lunch break between the second and third consultation, with the aim of in keeping the programme balanced, feasible and fun for the children. The results of this scoring were used to calculate the inter-rater reliability. After the first consultation, parents were asked to fill in an evaluation form on the acceptability and the burden of the scale.

### Composition of the final scale

Since all patients and parents evaluated the completeness and acceptability as 'good', no adaptations were made based on patient suggestions. Except for one boy with autism who was not able to complete only three out of four examinations during the pilot study because he was bored, all other children were able to complete the pilot study. Based on the experiences of the physicians involved and the inter-rater reliability, the pilot scoring list was adapted (see Supplementary Document 11). In addition, the instructions within the manual were clarified, especially for the physical examination domain. Subsequently, the list was again applied to two children with a mitochondrial disease with a mild and a severe phenotype who were not involved in the pilot study. In addition, the list was tested in patients in three other centres, who checked for acceptability and practical issues. After that, the scoring list was adapted again and sent to all participating centres. Again, adaptations were made in a Delphi-based process to compose the final scoring list, named the IPMDS (Supplementary Document 8 and 9). A 5-minute instructional video was prepared to illustrate the execution of some of the items (Supplementary Document 10).

### Testing the final scale

Feasibility, validity and reliability studies were performed in two to four children each using the IPMDS in the Departments of Paediatrics from the Universities of Pretoria, Hong Kong, München, Rotterdam, and the Children's Hospital of Philadelphia.. Children

with a mitochondrial disease, either based on pathological mutations in mtDNA or nuclear DNA or on mitochondrial dysfunction in muscle as measured by biochemistry, were eligible for inclusion. Exclusion criterion included treating physician expectation that travelling to the hospital would be too burdensome to the patient. Each individual centre selected children randomly. The number of patients and physicians was based on the local feasibility of executing the study.

## Feasibility

Feasibility was tested by asking (patients and) parents about their experiences after the first consultation and by counting the number of children in which the whole scoring list could be completed safely. The response was used as a measure for the raters' feasibility.

## Factor analysis

We used exploratory factor analysis to identify the underlying dimensions in the questionnaire.

## Construct validity

To test the construct validity of the factors, the hypothesis was proposed that patients rated as having severe disease by the physicians had higher sum scores for the factors compared to patients rated as having mild disease. Since there is no gold standard for mitochondrial disease severity, we used anchors to test construct validity. The NPMDS, a measure for global mitochondrial disease severity, was performed at the second consultation in every patient. The total score of the NPMDS was correlated to the total IPMDS score. In addition, all physicians were asked to rate: i) the general severity of the mitochondrial disease; ii) the subjective severity of the mitochondrial disease as experienced by parents; iii) the abnormalities at the physical examination; iv) the functional capabilities of the child, from 0 (not severe at all) to 10 (extremely severe) after every consultation. The PEDI was used to assess functional performance and abilities of the child. We hypothesised good correlation coefficients between: the total score of the NPMDS and the total IPMDS score; the general severity of the mitochondrial disease rated by physicians and the total IPMDS score; the subjective severity of the mitochondrial disease as experienced by parents and the IPMDS score; the abnormalities at the physical examination rated by physicians and Domain 2; and the functional capabilities of the child rated by physicians and Domain 3; the PEDI and Domain 3.

## Reliability

Inter-rater reliability was calculated using the scores of the same patient between physicians. Intra-rater reliability was tested by re-scoring the video recording of their own interview and examinations, approximately 6 months later (2 physicians from ErasmusMC, Rotterdam). Test-retest reliability was tested by asking parents to rate the

items within the first domain by telephone, both by the same rater also participating in the study (the University of Hong Kong) and between the mean score at the initial evaluation and evaluation by a nurse specialist (ErasmusMC, Rotterdam) one week later.

## Statistics

Because of the relatively small number of subjects included in our study, we used non-parametric tests for our analyses and reported median and range. The experiences of the parents and patients were assessed for each patient individually. From a rater perspective, the number of blank items was counted to test feasibility.

Factor analysis was used to explain the variance-covariance matrix for data in terms of relationships between a much smaller number of unobserved variables, called factors. We used the rater's data with the least missing data, missing items were replaced by the mean of the other three raters. We removed the items of which less than 80% was completed, including the items which could only be assessed in children older than 6 years old (hopping, running and rotating) and in case the child was not able to report complaints such as headache, gastroesophageal reflux or muscle pain or vibration or subtle touch at physical examination. Items with little variance (less than two items score 1 or more) were also removed for the factor analysis. Principal component analysis (PCA) was used as the extraction method for factors. The orthogonal rotation (varimax) with Kaiser's normalisation was used to simplify the interpretation of factors. The number of factors extracted was based on Eigen values  $>1$ . The adequacy of the sampling was determined using the Kaiser-Meyer-Olkin (KMO) measure. In addition, Bartlett's test of sphericity was applied to test if correlations between items were sufficiently large for PCA. The percentage of variance explained by each factor is also presented. Sum scores for the factors were calculated using clinically suitable items. Cronbach's alpha was calculated as a measure of internal consistency of each constructed factor. The difference between the patients with mild and with severe disease (the median value rated by physicians) for these sum scores was tested using a Mann-Whitney U test.

We used hypothesis testing to assess construct validity. Since mentioned hypotheses in this study involve two measurements of the same construct, we aimed at moderate to good correlations ( $\rho = 0.4$  to  $0.79$ ). We used Spearman's correlation coefficients to correlate between continuous or interval variables (IPMDS and its sub domains and functional parameters). The mean raters' score in each patient was used to calculate the correlation.

Inter-rater reliability between the physicians within one centre seeing one patient was calculated using intraclass correlation coefficient for agreement between raters ( $ICC_{\text{agreement}}$ ). Intra-rater reliability within two physicians was calculated using intraclass correlation coefficient for agreement between the rater's scores ( $ICC_{\text{agreement}}$ ). Test-retest reli-

bility was calculated using intraclass correlation coefficient for agreement between the scores ( $ICC_{\text{agreement}}$ ). An  $ICC_{\text{agreement}} \geq 0.7$  was used as 'acceptable'.<sup>19</sup>

A  $p$ -value of 0.05 was set for statistical significance. Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website (<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>). All analyses were performed using IBM's SPSS Statistics 22.0.0.1.

## Results

### Cohort description

A clinically and genetically heterogeneous cohort of seventeen children, aged 1.6 to 16 years, from five expert centres participated in this study (Supplementary Table 12). Rater details are presented in Supplementary Document 13.

### Scale description

After our Delphi-based process, the IPMDS consists of 61 items in three Domains: 23 in Domain 1 (Subjective complaints and symptoms; obtained by interviewing parents); 25 in Domain 2 (Physical examination; obtained by physical examination) and 13 in Domain 3 (Functional assessment; obtained by physical/motor function evaluation).

### Feasibility

The average time to complete the IPMDS was 35 minutes. Ninety-six percent of all items, relevant to the patient based on age and/or mental capacities, was completed (range 85 - 100%).

Sixteen (patients and their) parents filled out the feasibility questionnaire. All patients and parents indicated that the number of questions, the burden of the physical examination to the child, the duration of the interview, physical examination and the total time-burden were just right or could be (much) longer. The parents of a two-year-old patient experienced difficulties in translating the questions to the situation of a toddler. This was not reflected in the number of missing items for the toddlers (94%; range 91 – 97%) and in the inter-rater reliability for the first and the third domain ( $ICC_{\text{agreement}}$  0.86 versus 0.74 and 0.96 versus 0.97, respectively for toddlers versus all children). The inter-rater reliability of the second domain was low in toddlers compared to the whole cohort ( $ICC_{\text{agreement}}$  0.39 versus 0.81).

## Factor analysis

A total of 44 out of 61 items were included in the factor analysis. Based on factor loadings and communalities, we were able to include 32 out of the 49 items in one of the factors (Table 1A). Items with little variance were removed (items: 1.17. Diarrhoea, 2.03 Alertness, 2.04 Breathing, and 2.09 Nystagmus). Sample adequacy of the factor analysis was sufficient for all items within the IPMDS (KMO 0.51). Bartlett's tests of sphericity indicated that correlation between items were sufficiently large for PCA ( $p < 0.001$ ).

**Table 1.**

**Factor analysis.** A) The factor loadings for the individual items. Based on clinical communalities between items we composed three factors (in grey) including 32 out of the 49 included items; B) the factors and their characteristics resulting from factor analysis on the items within the IPMDS; C) Spearman's correlation coefficients between the sum scores of factors; D) median and IQR of the factor sum scores for patient groups within different disease stages and of the difference between these groups.

### A

	Basic functioning	Eating and digesting	Abnormalities at neurological examination
3.05 Sitting	0.926		
3.11 Reaching	0.916		
3.12 Grasping	0.901		
1.13 Communicating	0.871		
3.07 Standing	0.875		
3.04 Sitting up	0.867		
2.18 Hypertonia	0.848		
3.08 Walking	0.847		
3.06 Standing up	0.843		
3.02 Head control	0.831		
3.03 Rolling over	0.821		
2.20 Rigidity	0.808		
3.01 Communication	0.772		0.513
1.22a Continence day	0.592		
1.18 Cognitive	0.547		0.486
2.11 Prox muscle	0.408	0.614	
1.11 Swallowing		0.894	
1.10 Chewing	0.536	0.764	
1.14 Vomiting		0.74	0.476
2.01 Growth		0.507	-0.405
1.16 Constipation		-0.49	
2.02 Weight	-0.404	0.448	
2.08 Eye movements			-0.698

2.16 Tremor			-0.628
2.13 Hypokinesia	0.745		0.568
1.01 Response			0.543
2.19 Hypotonia			0.525
2.15 Ataxia			-0.511
1.23a Strabismus			-0.498
2.14 Abnormal movement	0.69		-0.496
1.12 Hearing	0.403		0.494
2.17 Reflexes			0.419
1.22b Continence night		0.75	
2.07 Strabismus		0.741	
2.10 Vision		0.717	
1.02 Exercise			0.616
1.09 Infections			
1.04 Tired			
1.23b Ptosis			
1.06 Epilepsy			
1.21 Breathing			
2.06 Ptosis			

Only factor loadings > 0.4 are presented

## B

Factor	<i>n</i>	Eigenvalues	% of explained variance	% of explained variance (total)	Cronbach's alpha
Basic functioning	16	14.3	31.1	31.1	0.97
Eating and digesting	6	5.2	14.1	45.2	0.76
Abnormalities at neurological examination	10	4.8	12.7	57.9	0.16

*n* = number of items in the factor

## C

	Basic functioning	Eating and digesting	Abnormalities at neurological examination
Basic functioning	1.00		
Eating and digesting	0.47**	1.00	
Abnormalities at neurological examination	0.42**	0.13	1.00

\* significant at the 0.05 level; \*\* significant at a 0.01 level.



## D

Factor	Mild disease ( <i>n</i> = 3)		Severe disease ( <i>n</i> = 10)		Difference between mild and severe patients
	median	IQR	median	IQR	
Basic functioning	0	(0 - 0)	43	(14 - 56)	0.01
Eating and digesting	2	(0 - 2)	10	(5 - 16)	0.08
Abnormalities at neurological examination	3	(1 - 3)	10	(2 - 12)	0.29

*IQR: interquartile range*

We identified three factors (Table 1A and 1B). Based on the items included in these factors, they were named: 1: basic functioning (31.1% of explained variance), 2: eating and digesting (14.1% of the explained variance) and 3: abnormalities at neurological examination (12.7% of explained variance). The internal consistency of these factors was good for factor 1 (Cronbach's alpha = 0.97), acceptable for factor 2 (Cronbach's alpha = 0.76) and poor for factor 3 (Cronbach's alpha = 0.16). The total percentage of explained variance by the three factors was 57.9%. The Spearman's correlation coefficient's between the factors (Table 1C), indicate that all factors represent a unique identity.

### Construct validity

We hypothesised that patients with a severe general mitochondrial disease severity would show higher median scores on the extracted factors compared to patients with mild disease (Table 1D). This hypothesis was confirmed for factor 1(basic functioning;  $p = 0.01$ ;  $n = 13$ ), but not for factor 2 (eating and digesting) and 3 (abnormalities at neurological examination ( $p = 0.07$  and  $0.28$ , respectively). In Table 2 the correlation coefficients between the factors, the IPMDS<sub>total</sub>, the IPMDS sub domains, the severity scores, the PEDI and the NPMDS are presented. The correlation between the basic functioning factor and the PEDI, rating the functional performance and abilities of the child, was excellent ( $\rho = 0.90$ ;  $p < 0.001$ ); the correlation between the abnormalities at neurological examination factor and the rater-rated severity of abnormalities at general physical examination was weak ( $\rho = 0.23$ ;  $p < 0.05$ ). All pre-defined hypotheses used for construct validity testing indeed had good to excellent correlations ( $\rho = 0.64 - 0.90$ ; Table 2).

**Table 2.**  
**Construct validity of the IPMDS.**  
Correlation coefficients of the total score and sub-scores of the IPMDS with subjective disease severity as rated by parents, subjective disease severity by the physicians (mean), with the NPMDS total and sub domain scores and with the PEDI total and sub domains scores.

	IPMDS - total	IPMDS - Domain 1	IPMDS - Domain 2	IPMDS - Domain 3	Sum of basic functioning	Sum score of eating and digesting	Sum score of abnormalities at neurological examination
General disease severity (P)	<b>0.806**</b>	0.480**	0.558**	0.725**	0.658**	0.426*	0.038
General disease severity (R)	<b>0.642**</b>	0.529**	0.744**	0.810**	0.786**	0.476**	0.237*
Severity of the abnormalities at physical examination (R)	0.853**	0.536**	<b>0.750**</b>	0.839**	0.803**	0.519**	<b>0.226*</b>
Functional limitations (P)	0.665**	0.500**	0.563**	<b>0.779**</b>	<b>0.740**</b>	0.0328	0.135
Functional limitations (R)	0.817**	0.573**	0.719**	<b>0.809**</b>	<b>0.760**</b>	0.381**	0.293**
PEDI - total	0.912**	-0.784**	-0.773**	<b>-0.890**</b>	<b>-0.904**</b>	-0.668**	-0.283
PEDI - self-care	-0.916**	-0.767**	-0.767**	-0.881**	-0.882**	-0.684**	-0.263
PEDI - mobility	-0.918**	-0.749**	-0.797**	-0.917**	-0.901**	-0.588**	-0.243
PEDI - social function	-0.673**	-0.750**	-0.569**	-0.662**	-0.713**	-0.552**	-0.297
NPMDS - total	<b>0.840**</b>	0.475**	0.771**	0.836**	0.816**	0.572**	0.183
NPMDS - Section 1	0.887**	0.512**	0.777**	0.865**	0.882**	0.606**	0.27
NPMDS - Section 2	-0.0144	-0.172	-0.117	0.100	0.029	-0.0133	-0.146
NPMDS - Section 3	0.854**	0.597**	0.855**	0.773**	0.657**	0.676**	0.196
IPMDS - total	1.000	0.836**	0.929**	0.907**	<b>0.884**</b>	<b>0.557**</b>	<b>0.553**</b>
IPMDS - Domain 1	0.836**	1.000	0.551**	0.551**	0.750**	0.632**	0.439**
IPMDS - Domain 2	0.929**	0.551**	1.000	0.883**	0.836**	0.615**	<b>0.529**</b>
IPMDS - Domain 3	0.907**	0.551**	0.883**	1.000	0.931**	0.528**	0.426**

\* significant at the 0.05 level; \*\* significant at a 0.01 level; in bold the correlations used for the construct validity hypothesis testing.

## Reliability

Table 3 and Supplementary Table 14 and 15 show the inter-rater, intra-rater and test-retest reliability of the IPMDS. The inter-rater reliability of the sum scores of the three factors was good ( $ICC_{\text{agreement}} = 0.98, 0.94, \text{ and } 0.78$  for basic functioning, eating and digesting, and abnormalities at neurological examination, respectively). The median  $ICC_{\text{agreement}}$  of all individual items within the IPMDS was 0.85 (range 0.23 – 0.99). The median  $ICC_{\text{agreement}}$  was 0.81 (range 0.44 – 0.98) of the first domain, 0.74 (range 0.23 – 0.93) of the second domain and 0.97 (range 0.93 – 0.99) in the third domain (Table 3). Video rated intra-rater reliability was good (median  $ICC_{\text{agreement}} = 0.87$ ; Supplementary Table 14). Test-retest reliability by telephone interviewing of the first Domain after one week was excellent when performed by the same rater (median  $ICC_{\text{agreement}} = 1.0$  in 15 out of 16 items with variance) but inconsistent when performed by a different rater (median  $ICC_{\text{agreement}} 0.73$ ; Supplementary Document 16).

**Table 3.**  
**Inter-rater reliability.**

Inter-rater reliability ( $ICC_{\text{agreement}}$ ) per item of the IPMDS. Items with an  $ICC_{\text{agreement}} \geq 0.7$  (good agreement) are marked in green;  $\leq 0.3$  (poor agreement) marked in blue.

		ICC-agreement	n =	
1.01	Response to environment	0.670	17	
1.02	Exercise capacity	0.701	13	x
1.03	Walking distance	0.847	16	
1.04	Tiredness	0.780	16	
1.05	Depressed	0.847	12	x
1.06	Epilepsy	0.855	17	
1.07	Headache	0.931	12	x
1.08	Muscle pain	0.575	12	x
1.09	Infections	0.581	17	
1.10	Chewing	0.942	17	
1.11	Swallowing	0.886	17	
1.12	Hearing	0.762	17	
1.13	Communicating	0.953	16	
1.14	Vomiting	0.972	17	
1.15	Gastroesophageal reflux	0.962	15	
1.16	Constipation	0.833	17	
1.17	Diarrhoea	0.976	16	
1.18	Cognitive functioning	0.635	15	
1.19	Behavioural problems	0.443	16	
1.20	Autistic features	0.661	11	x
1.21	Breathing subjective	0.934	15	
1.22a	Continence day	0.725	14	
1.22b	Continence night	0.725	14	

1.23a	Strabismus when tired	0.736	15	
1.23b	Ptosis when tired	0.573	15	
1.23c	Dysarthria when tired	0.838	13	x

x: feasible in less than 80% of study participants

		ICC-agreement	n =	
2.01	Growth	0.944	15	
2.02	Weight	0.751	15	
2.03	Alertness	NV	17	
2.04	Breathing	0.740	17	
2.05	Dysarthria	0.901	12	x
2.06	Ptosis	0.553	17	
2.07	Strabismus	0.82	16	
2.08	Eye movements	0.691	16	
2.09	Nystagmus	0.925	16	
2.10	Vision	0.343	11	x
2.11	Proximal muscle power	0.913	15	
2.12	Distal muscle power	0.738	14	
2.13	Hypokinesia	0.306	14	
2.14	Abnormal movements	0.884	16	
2.15	Ataxia	0.858	14	
2.16	Tremor	0.864	15	
2.17	Reflexes	0.312	16	
2.18	Hypertonia	0.649	17	
2.19	Hypotonia	0.470	17	
2.20	Rigidity	0.597	17	
2.21a	Vibration	0.231	11	x
2.21b	Subtle touch	NV	11	x

x: feasible in less than 80% of study participants NV = no variance

		ICC-agreement	n =	
3.01	Communicating	0.957	17	
3.02	Head control	0.983	17	
3.03	Rolling over	0.985	16	
3.04	Sitting up	0.963	17	
3.05	Sitting up	0.984	17	
3.06	Standing up	0.979	17	
3.07	Standing	0.974	17	
3.08	Walking	0.989	17	
3.09	Running	0.968	13	only >6 years
3.10	Hopping	0.954	13	only >6 years
3.11	Reaching	0.964	16	
3.12	Grasping	0.949	16	
3.13	Rotating	0.985	14	only >6 years

## Discussion

The International Paediatric Mitochondrial Disease Scale (IPMDS), a multi-dimensional scale rating clinically relevant aspects of mitochondrial disease in children, was developed during a Delphi-based process by an international expert team in close dialogue with patients and their parents. Designing a scale for general mitochondrial disease severity is challenging because of the wide variability in symptoms and (dis)abilities seen in children with mitochondrial disease. This complexity is reflected in the high number of items included in the IPMDS and the breadth of the items reflecting the multi-systemic nature of the disease. Critically evaluating the psychometric properties in seventeen children in five international expert centres, we found good feasibility and an acceptable construct validity and reliability.

We found a suboptimal inter-rater reliability for some of the items of the IPMDS, mainly within the physical examination domain of the IPMDS. This is in agreement with literature, where low inter-rater agreement in gait analysis, neurological reflexes and classification of movement disorders has been reported.<sup>20-23</sup> Lower agreement was observed in patients with developmental delay<sup>24</sup> and in general neurologists compared to residents and experts.<sup>20</sup> Our teams of raters indeed reported difficulties in agreeing on the hypertonia and rigidity. This might be explained by the mixed picture of pyramidal and extrapyramidal syndromes in patients with mitochondrial disease (and more specifically in patients with Leigh syndrome; about half of the cases included in this study). Moreover, tone is highly dependent on the child's level of alertness, emotional state, fatigue and posture<sup>25</sup> and changes in tone during the day were frequently reported by our raters. One option is to remove Domain 2 from our scale. This option will be studied in more detail when we have obtained more experience and patient data on the IPMDS.

Although the IPMDS was adopted from the NPMDS, there are important differences. First of all, the NPMDS consists of 26 items with mostly a 0-3 scale, whereas the IPMDS consists of 61 items, mostly on a 0-5 scale. For example, the 'feeding' item in the NPMDS was replaced by items on chewing (including the ability to chew e.g. bread crusts and meat), swallowing (including difficulties/choking with dry food or fluids) and vomiting. The IPMDS therefore takes longer to execute (on average: 35 minutes), but is considerably more detailed. Although the burden to the patient doesn't seem to be too high, as indicated by the positive evaluation of the IPMDS by patients and parents, the feasibility of performing the IPMDS as part of daily care at the outpatient clinic, a central tenet of developing the NPMDS, is questionable. Secondly, since behavioural problems and speech and language problems were among the most burdensome complaints experienced by patients and their parents,<sup>11</sup> we included these items in the complaints and symptoms and in the functional (communication) domain of the IPMDS. Thirdly, the IPMDS has a functional domain that is expected to be the most objective, responsive

and relevant item for natural history and intervention studies. Lastly, we use the same scoring system for all ages (with the exception of some functional items). Although this complicates the analysis in very young children – which was also difficult in the NPMDS 0-24 months since newborns differ from toddlers – longitudinal analysis is more meaningful when using the same scoring system. These differences are illustrated by the lack of a statistically significant correlation between the IPMDS and the NPMDS.

Strengths of our study include: following of the applicable guidelines provided by the US Food and Drug administration<sup>26</sup> for the design of the IPMDS, the international expert team giving input to the content of the IPMDS and the validation in five international expert centres in a heterogeneous group of children with mitochondrial disease. However, since we aimed at a scale covering the complete phenotypic spectrum of mitochondrial diseases, the IPMDS may contain many irrelevant items for individual patients. Other weaknesses include the undetermined influence of normal development on the score and relatively small study population per centre. The small population also affects the validity of the factor analysis and exploratory factor analysis should be repeated when more data is obtained.

Based on our data, we suggest using the IPMDS for the assessment of natural history in children with mitochondrial diseases. Since the IPMDS also includes subjective and functional parameters, this natural history data will ultimately not only provide clinicians with relevant information for clinicians about which patients will be at risk to develop e.g. a cardiomyopathy or renal failure, but will also provide relevant prognostic information to patients and their parents. The data collected in these natural history studies should be used for another exploratory factor analysis to obtain further evidence for the construct validity of the IPMDS (we kindly invite you to upload your IPMDS scoring sheets on our website <http://www.ncmd.nl/ipmids>). Besides, this data will facilitate setting age-limits for the IPMDS, since both the number of missing items and the inter-rater reliability indicate that adaptations necessary for babies and toddlers. In parallel, the minimal clinically important difference should be assessed, for example by assessing which changes patients and parents feel to be relevant compared to the previous assessment in a prospective natural history study or by statistically-based approach, although the latter does not take into account the patients' perspective.<sup>27</sup>

In conclusion, the IPMDS seems a robust tool for the follow-up of children with mitochondrial diseases. The data obtained in natural history studies in combination with a close dialogue with parents regarding the minimal clinically important difference will further substantiate the instrument in the near future.

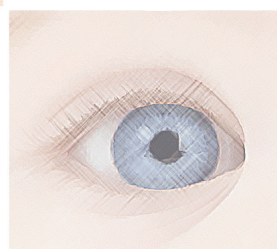
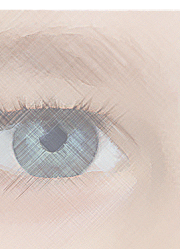
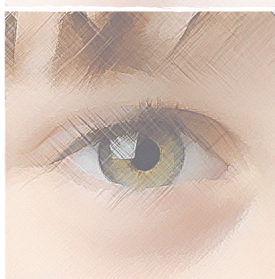
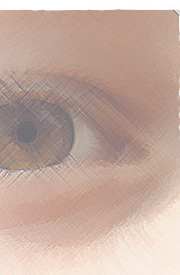
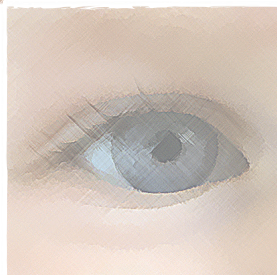
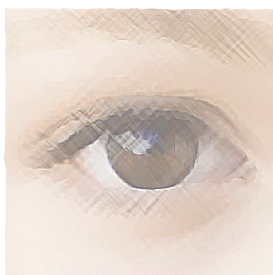
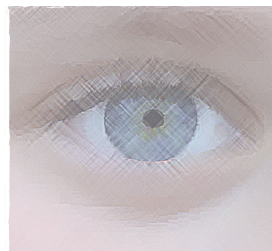
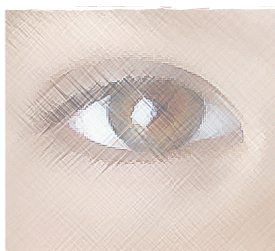
## References

1. Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders--past, present and future. *Biochimica et biophysica acta* 2004;1659:115-20.
2. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med* 2012;366:1132-41.
3. Koene S, Smeitink J. Mitochondrial medicine: entering the era of treatment. *Journal of internal medicine* 2009;265:193-209.
4. Viscomi C, Bottani E, Civiletto G, *et al.* In vivo correction of COX deficiency by activation of the AMPK/PGC-1alpha axis. *Cell metabolism* 2011;14:80-90.
5. Wenz T. PGC-1alpha activation as a therapeutic approach in mitochondrial disease. *IUBMB Life* 2009;61:1051-62.
6. Viscomi C, Bottani E, Zeviani M. Emerging concepts in the therapy of mitochondrial disease. *Biochimica et biophysica acta* 2015;1847:544-57.
7. Blanchet L, Smeitink JA, van Emst-de Vries SE, *et al.* Quantifying small molecule phenotypic effects using mitochondrial morpho-functional fingerprinting and machine learning. *Scientific reports* 2015;5:8035.
8. Peng M, Ostrovsky J, Kwon YJ, *et al.* Inhibiting cytosolic translation and autophagy improves health in mitochondrial disease. *Human molecular genetics* 2015;24:4829-47.
9. Koene S, Jansen M, Verhaak CM, De Vruet RL, De Groot IJ, Smeitink JA. Towards the harmonization of outcome measures in children with mitochondrial disorders. *Developmental medicine and child neurology* 2013.
10. Phoenix C, Schaefer AM, Elson JL, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord* 2006;16:814-20.
11. Koene S, Wortmann SB, de Vries MC, *et al.* Developing outcome measures for pediatric mitochondrial disorders: which complaints and limitations are most burdensome to patients and their parents? *Mitochondrion* 2013;13:15-24.
12. Beenakker EA, Maurits NM, Fock JM, Brouwer OF, van der Hoeven JH. Functional ability and muscle force in healthy children and ambulant Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2005;9:387-93.
13. Parreira SL, Resende MB, Zanoteli E, Carvalho MS, Marie SK, Reed UC. Comparison of motor strength and function in patients with Duchenne muscular dystrophy with or without steroid therapy. *Arq Neuropsiquiatr* 2010;68:683-8.
14. Abel MF, Damiano DL, Blanco JS, *et al.* Relationships among musculoskeletal impairments and functional health status in ambulatory cerebral palsy. *J Pediatr Orthop* 2003;23:535-41.
15. Schaefer AM, Phoenix C, Elson JL, McFarland R, Chinnery PF, Turnbull DM. Mitochondrial disease in adults: a scale to monitor progression and treatment. *Neurology* 2006;66:1932-4.
16. Feldman AB, Haley SM, Coryell J. Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. *Phys Ther* 1990;70:602-10.
17. Van Dillen LR, Roach KE. Interrater reliability of a clinical scale of rigidity. *Phys Ther* 1988;68:1679-81.
18. Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord* 2005;15:463-70.
19. Vet HTCM, LB; Knol, DL. Measurement in Medicine. Cambridge: Cambridge University Press; 2011.

20. Beghi E, Regio V, Papantonio A, *et al.* Reliability of clinical diagnosis of dystonia. *Neuroepidemiology* 2014;43:213-9.
21. Miller TM, Johnston SC. Should the Babinski sign be part of the routine neurologic examination? *Neurology* 2005;65:1165-8.
22. Singerman J, Lee L. Consistency of the Babinski reflex and its variants. *Eur J Neurol* 2008;15:960-4.
23. Bella GP, Rodrigues NB, Valenciano PJ, Silva LM, Souza RC. Correlation among the Visual Gait Assessment Scale, Edinburgh Visual Gait Scale and Observational Gait Scale in children with spastic diplegic cerebral palsy. *Revista brasileira de fisioterapia* 2012;16:134-40.
24. Hsue BJ, Chen YJ, Wang YE. The intra- and inter-rater reliability of component analysis of rise from supine in the children with typical development and developmental delay. *Res Dev Disabil* 2014;35:162-70.
25. Sanger TD, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Classification and definition of disorders causing hypertonia in childhood. *Pediatrics* 2003;111:e89-97.
26. Administration FaD. FDA Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
27. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *The Journal of manual & manipulative therapy* 2012;20:160-6.







# Chapter 8

The assisted 6-minute  
cycling test: an exploratory  
study in children



## The assisted 6-minute cycling test: an exploratory study in children

Ilse Dirks<sup>1</sup>, **Saskia Koene**<sup>1</sup>, Renee Verbruggen<sup>2</sup>, Jan A.M. Smeitink<sup>1</sup>, Merel Jansen<sup>2</sup>,  
Imelda J.M. de Groot<sup>1,2</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders at the Department of Paediatrics, Radboudumc, Nijmegen, The Netherlands

<sup>2</sup>Donders Centre for Neuroscience, Department of Rehabilitation, Radboudumc, Nijmegen, The Netherlands

Muscle & Nerve, accepted manuscript online 30 Dec 2015

The 6-minute walking test is frequently used as an outcome measure for clinical trials in neuromuscular disease. Since this submaximal endurance test is not feasible for non-ambulatory patients, the motor-assisted 6-minute cycling test was developed.

Nineteen children with neuromuscular disorders and children with OXPHOS dysfunction performed the cycling test and the 6-minute walking test to explore feasibility and construct validity. Test-retest reproducibility was tested within three weeks.

The assisted 6-minute cycling test was feasible in 90% and 78% of the patients with a neuromuscular disorder and OXPHOS-dysfunction, respectively. Heart rate indicated the submaximal character of the test in the majority of patients. The assisted 6-minute cycling test for legs correlated with the 6-minute walking test in patients with OXPHOS dysfunction, but did not reach significance in patients with neuromuscular disease. Acceptable reproducibility was shown for both legs and arms.

This exploratory study indicates that the assisted 6-minute cycling test seems a promising outcome measure for specific subgroups of patients with a neuromuscular disorder and patients with OXPHOS-dysfunction. Before we can recommend the 6-minute cycling test over the 6-minute walking test more studies, especially in weaker and/or non-ambulatory children, are indicated.

## Introduction

More and more therapeutic studies are being performed in children with neuromuscular disorders and patients with mitochondrial oxidative phosphorylation dysfunction in fresh muscle specimen (OXPHOS dysfunction). To measure the effect of these promising interventions, it is important to have reliable, valid and responsive outcome measures that measure clinically relevant aspects of the child's abilities and limitations. Since endurance and fatigability are important determinants of the patients' abilities to perform the activities of daily life,<sup>1</sup> it seems reasonable to use instruments measuring these disease aspects in clinical trials.

Endurance can be measured using maximal and submaximal endurance tests. The submaximal test only gives an indication of the maximal endurance.<sup>2</sup> Children with a neuromuscular disorder are mainly limited in performing physical activities requiring endurance of muscle function rather than they are limited in their maximal aerobic power.<sup>3</sup> Therefore, submaximal exercise tests are more relevant to and more feasible for these children than maximal exercise tests. Testing submaximal endurance provides data on the endurance required to perform the activities of daily life<sup>1</sup> and previous studies showed promising results in patients with Duchenne muscular dystrophy (DMD).<sup>4</sup> Currently, there is no gold standard to measure submaximal endurance in children. The most commonly used test to measure submaximal endurance in neuromuscular disorders is the 6-minute walk test (6MWT).<sup>1</sup> However, this test could only be used in patients who are able to walk and are not expected to lose the ability to walk during the trial.

The assisted 6-minute cycling test (A6MCT) for legs and arms was developed as a non-invasive and quantitative method to measure submaximal endurance in ambulatory and non-ambulatory children with a neuromuscular disorder. A previous study by our group showed that the A6MCT was feasible in >90% of the boys with DMD and that the distance cycled correlated with the distance walked at the 6MWT.<sup>5</sup> It is likely that this test is not specific for DMD only, but also applicable to other neuromuscular disorders, including OXPHOS dysfunction. However, other factors could play a role such as behaviour problems, present in both some types of neuromuscular disorders and OXPHOS dysfunction, or the difference between the presence of paresis at start (neuromuscular) or developing/increasing paresis during the test (OXPHOS dysfunction).

The aim of this exploratory study was to evaluate the psychometric properties (feasibility, validity and test-retest reproducibility) of the A6MCT in a group of children with neuromuscular disorders other than DMD and in children with OXPHOS dysfunction.

## Methods

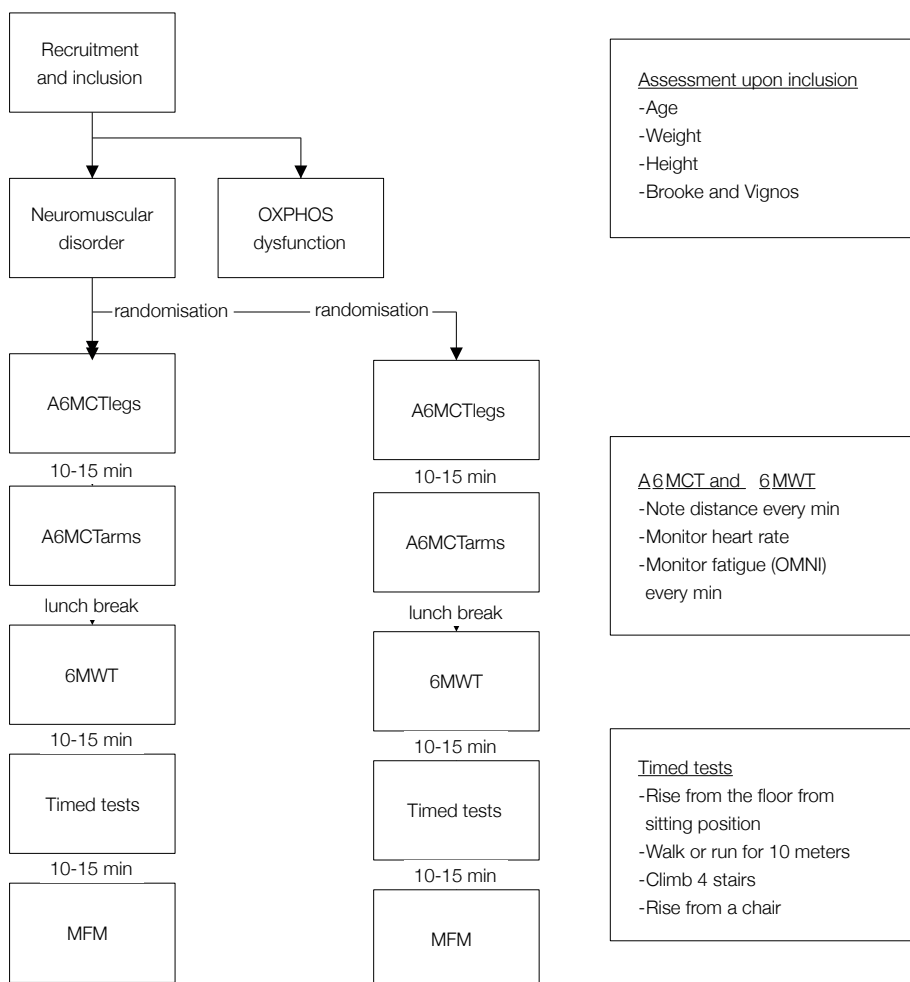
### Study population

Children with neuromuscular disorders and children with OXPHOS dysfunction in muscle, regularly seen at the Radboudumc, Nijmegen, The Netherlands (Department of Rehabilitation and the Nijmegen Centre for Mitochondrial Disorders, respectively) were selected for this study. Children aged 6 - 18 years with a confirmed neuromuscular diagnosis of myotonic dystrophy (MD), spinal muscular atrophy type 3 (SMA), limb-girdle muscular dystrophy (LGMD), or congenital myopathy (CM) were included. Children with OXPHOS dysfunction were included when mitochondrial dysfunction was detected in fresh muscle according to one or more of the following criteria: i) decrease in ATP production ( $\leq 90\%$  of lowest reference value); and/or ii) deficiency of one or more enzyme complexes of the oxidative phosphorylation system (below the lowest reference value) and/or iii) a pathogenetic mutation in mitochondrial or nuclear DNA. Both children with a neuromuscular and OXPHOS dysfunction needed to be able to walk for at least 75 meters. Exclusion criteria in both groups included severe behavioural problems, severe cognitive impairment, severe respiratory insufficiency, high likelihood of falls and/or symptomatic cardiomyopathy. Patients (and their parents) who met the inclusion criteria were informed about the study by their physician. Interested patients and parents were further informed about the study by an information letter in which they were invited to participate in the study. All parents, and participants who were over 12 years old and mentally competent, gave their written informed consent. This study has been conducted after approval from the Medical Research Ethics Committee Arnhem-Nijmegen (MREC NL.37246.091.11).

### Study protocol

All tests were performed at the outpatient clinic at the Department of Rehabilitation at the Radboudumc. All patients followed the same time-schedule, including a lunch break of 60 minutes and a pause of 10 - 15 minutes after each test in which patients could have a drink (see Figure 1). Patients were asked to wear comfortable clothes and to use a wheelchair to come to the outpatient clinic where possible. Height and weight were measured before the tests were performed.





**Figure 1.**

**Flowdiagram from recruitment to experiments in both children with a neuromuscular disorder and OXPHOS dysfunction.** A6MCT = assisted 6-minute cycling test; 6MWT = 6-minute walk test; MFM = motor function measure.

First the Brooke and Vignos scale were determined. The Brooke scale grades upper limb function on an ordinal scale from 1 - 6; 1 means that the patient is able to abduct the arms in a full circle until they touch above the head, and 6 means that patients are unable to raise their hands to their mouth and have no useful function of the hands. The Vignos scale grades lower limb function on an ordinal scale ranging from 1 - 10, with 1 meaning that the patient can walk and climb stairs without assistance and 10 means that the patient is bed-bound.<sup>6</sup>

Subsequently, patients with a neuromuscular disorder performed the A6MCT<sub>legs</sub> followed by the A6MCT<sub>arms</sub> (see Figure 2 for starting position of both tests). To exclude the influence of the commonly reported general fatigue after one A6MCT influencing the result of the second A6MCT, patients with OXPHOS dysfunction were randomly assigned to the 'legs first' or 'arms first' group. Both tests were performed according to the previously described protocol by Jansen *et al.*<sup>5</sup> The participants were asked to cycle as fast as possible and keep this up for 6 minutes. Every 15 seconds the children were encouraged with standardised phrases to maintain attention and to complete the test as well as possible. Positioning and motor assistance were standardised (passive mode 1, no-load speed 7 RPM). During both A6MCTs, the number of revolutions (rev.) was noted every minute, with the primary outcome being the total number of rev. in 6 minutes.



**Figure 2.**

**Initial position of the A6MCT<sub>legs</sub> and A6MCT<sub>arms</sub>.** This figure has been previously published in Jansen *et al.* (2010)<sup>7</sup>. Permission was obtained from the original publisher, BioMedCentral, to reproduce the figure.

After lunch, the 6MWT was performed, following the protocol of McDonald *et al.*,<sup>8</sup> except for that the parents were allowed to be present, but not to encourage the child. Patients were asked to walk (not run or hop) as fast as possible for 6 minutes on a 25 meter course. They were allowed to rest, but not to sit or lay down (timing continued during resting). During the 6MWT, one of the researchers walked behind the patient for safety reasons and to encourage the patient every 15 seconds with standardised phrases. The distance walked was recorded every minute, with the primary outcome being the distance walked in 6 minutes. During both A6MCTs and the 6MWT, fatigue was monitored every minute by measuring heart rate using a standard heart rate monitor (Onyx Classic; Sigma, Germany) and the OMNI scale for perceived exertion that grades the perception of fatigue on a scale from 0 (not tired at all) to 10 (very very tired) and contains both verbal and pictorial descriptions.<sup>9,10</sup>

After the 6MWT and a short break, the timed tests were performed to measure the functional abilities of the patients.<sup>6,11</sup> They were asked to rise from the floor from sitting position, walk or run for 10 meters, climb 4 stairs and rise from a chair. The time and the performance (the presence of compensatory movements, the use of help of hands on thighs, arm-rest, handrail or chair) was noted.<sup>12</sup> Since these tests are performed differently in various studies, these results are not reliable and therefore we cannot compare our results with other studies. Finally, the Motor Function Measure (MFM) was performed to assess motor performance.<sup>13,14</sup> This test is composed of 32 items in three dimensions: standing position and transfers, axial and proximal motor function, and distal motor function. The maximum score on this test is 100%, a higher score reflects better motor function.

After finishing the last test, parents received an evaluation form (designed by the authors for the purpose of this study) to report possible complications, such as an increase in fatigue, muscle pain, muscle weakness, epilepsy, coordination problems, behaviour problems, concentration problems or other problems in the day after the test. Items were scored on a 4 point scale: a lot more than normally (3), slightly more than normally (2), barely more than normally (1) and not different from normal (0).

### Feasibility

The A6MCT is rated as successful if the test has been completed (cycling the full 6 minutes, if necessary with pauses but restart within 15 seconds) without severe complications like muscle pain or stiffness the day after the test.

## Test results

The results of the A6MCT for boys were compared to the results of Jansen *et al.*<sup>5</sup> Girls were compared to the z-scores for healthy girls obtained in a pilot study in 50 healthy girls. Results of the 6MWT were compared with gender-, height- and age-matched controls.<sup>19</sup>

## Validity

First, we assessed whether the A6MCT measured submaximal exercise capacity by interpreting heart rate, subjective fatigue and the slope of the distance per minute curve. Submaximal exercise capacity was defined as 70% of maximal heart rate and the targeted OMNI score was defined as  $\leq 6$ .<sup>5</sup> Secondly, since there is no gold standard to measure endurance in patients with MD, SMA, LGMD, CM, and patients with OXPHOS dysfunction we used construct validity for the validation of the A6MCT. We used hypothesis testing to assess construct validity. The results of the A6MCTs were correlated to the results of the test rating walking capacity (6MWT), gross motor function (MFM) and functional abilities (timed tests) to determine construct validity. We hypothesised a good correlation between the A6MCT<sub>legs</sub> and the 6MWT and moderate correlations between the other motor function parameters and the A6MCTs.

## Reliability

Children were asked to participate in a retest after 4 weeks. All children who agreed to participate were eligible for inclusion.

## Statistical analysis

In this explorative study, data were analysed separately for the neuromuscular disorders and children with OXPHOS dysfunction in muscle. Since this study included a small number of patients per group continuous data were expressed as median and inter quartile range (IQR) and ordinal data were expressed as median with range. Construct validity was calculated by correlating the results of the different tests by Spearman's correlation coefficients. Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website.<sup>15</sup> The Wilcoxon signed-rank test was used to compare median scores in case of paired measurements. To evaluate test-retest reproducibility, intraclass coefficient ((ICC) absolute agreement) was calculated to compare the number of rev. on the A6MCT<sub>leg/arms</sub> and the distanced walked at the 6MWT at the test and at the retest. An  $\alpha$  of 0.05 was used for statistical significance. Because of the small numbers and the exploratory character of this study, we did not use Bonferroni adjustment. Statistical analyses were performed using SPSS (version 20.0) for Windows (SPSS, Inc. Chicago, Illinois).

## Results

### Patient characteristics

Nineteen children participated in this study. Ten of them (five boys, five girls) had a neuromuscular disorder, whereas nine children (eight boys, one girl) had mitochondrial OXPHOS dysfunction in muscle. Characteristics of the children are described in Table 1A and 1B.

### Feasibility of the tests

Both A6MCTs were feasible in nine out of ten patients with a neuromuscular disorder (90%) and in seven out of nine (78%) patients with a OXPHOS dysfunction. In one neuromuscular patient (ID 8, with MD) the A6MCTs were not feasible during the retest because of behavioural problems. The A6MCTs were not feasible in the two children with OXPHOS who had a mental retardation and failed to complete the tests because of behavioural problems (previously not rated as severe) and/or decreased attention span and lack of cooperation and understanding of the tasks. In OXPHOS dysfunction patient 1 and 8, both pubertal boys with near-normal muscle strength and coordination problems, the bicycle was unstable during the exercise. Therefore, both patients seemed to be limited in reaching their maximally possible number of rev.

The 6MWT was not feasible for three children (feasibility 70%, ID, 6, 8 and 9) with a neuromuscular disorder when following the McDonald protocol. One child with SMA (ID 6) had to take pauses after three minutes of walking and during the last minute she was not able to walk anymore. In a patient with MD (ID 8) the retest was not feasible because of behavioural problems. One child with a CM (ID 9) had to sit down after three minutes of walking. The 6MWT was not feasible in three children with OXPHOS dysfunction in muscle (feasibility 66%, ID 2, 3 and 5). Two patients with developmental delay (ID 2 and 5) were not able to complete the test because of decreased attention span and lack of cooperation and understanding of the tasks. In patient 3, the first 6MWT was not valid because he was joking around and not doing the best he could.

No serious adverse events were reported during or directly after performing the A6MCT and 6MWT. To report possible complications the day after the tests, parents were asked to fill in a complication form. Eleven complication forms from seven patients with a neuromuscular disorder (four patients completed the form for both the test and the retest) were returned to the researchers. The median scores were 0 (not more than normally) for tiredness (IQR 0 - 2; range 0 - 3) muscle weakness (IQR 0 - 1.25; range 0 - 2), balance (IQR 0 - 0; range 0 - 1), coordination problems (IQR 0 - 0; range 0 - 0), behavioural (IQR 0 - 0; range 0 - 2), and concentration problems (IQR 0 - 0; range 0 - 0) the day after the test was. The median score for muscle pain was 1 (barely more than normally; IQR 0 - 2; range 0 - 3).

ID	Gender	Genetic mutation	ATP production (% of lowest reference value)	Enzyme complex deficiencies	Age (years)	Develop-mental age (years)	Weight (kg)	Height (cm)	BMI (kg/m2)	Wheelchair for long distances?	A6MCT arms (rev.)	Z-scores A6MCT arms	Retest A6MCT arms	A6MCT legs (rev.)	Z-scores A6MCT legs	Retest 6MCT legs	6MWT (m)	Z-scores 6MWT	MFM (%)	Brooke	Vignos
1	M	Pending	32	CIII	13	6	67	168	24.8	No	823	-0.57	846	684	-1.39	775	625	-1.2	93	1	2
2§	M	SDHA1	86	CII	9	3.4	28	129	16.8	Yes	NF	NF		NF	NF		NF	NF	NF	1	7
3#	M	Pending	84	Normal	8	8	33	131	19.0	Yes	499	-2.11	634	795	-0.14	801	NF	NF	NF	1	1
4	M	SEPY1	50	Normal	8	8	21	120	16.7	Yes	542	-1.72	560	595	-3.15	700	312	-4.1	91	1	2
5§	M	Pending	27	MECD	7	2.3	24	131	14.0	No	NF	NF		NF	NF		NF	NF	NF	NF	NF
6	F	Pending	50	MECD	6	6	31	129	18.6	Yes	581	-0.24		623	-0.64		421	-2.1	94	1	1
7	M	Pending	40	Normal	9	9	34	143	16.6	Yes	638	-1.13	668	599	-2.79	646	357	-4.2	97	1	1
8	M	Pending	80	Normal	10	10	47	150	20.9	Yes	600	-1.75		665	-1.67		573	-1.3	97	NA	NA
9	M	Pending	90	Normal	9	9	34	143	16.6	Yes	748	-0.14	628	800	-0.34	854	447	-2.8	99	1	1
Median	-	-	50	-	9.05	-	33	131	16.8	-	600	Mean -1.10	634	665	Mean -1.44	775	434	Mean -2.6	94		
(IQR)			(36-59)		(7.5-9.5)		(26-41)	(129-147)	(16.6-20.0)		(542-748)		(594-757)	(599-795)		(673-827)	(346-596)		(91-97)		

\* = Abnormal gait pattern; † = Using hands; ‡ = Abnormal; § = Both the A6MCTs and the 6MWT not feasible because of behavioural problems and attention deficit; # = 6MWT not feasible because of behavioural problems; 6MWT = 6-minute walk test; A6MCT = Assisted 6-minute cycling test; BMI = Body Mass Index; cm = centimetres; C = Complex; kg = Kilogram; m = meters; MECD = Multiple Enzyme Complex Deficiencies; MFM = Motor Function Measure; MRC = Medical Research Council; NA = not available; NF = not feasible; Rev = revolutions; sec = seconds;

**Table 1.**  
**Clinical characteristics of children with neuromuscular disorders (A) and OXPHOS dysfunction (B).**

**A**

ID	Gender	Neuromuscular disorder	Age (years)	Develop-mental age (years)	Weight (kg)	Height (cm)	BMI (kg/m2)	MRCarms	MRClegs	Wheelchair for long distances	A6MCT arms (rev.)	Z-scores A6MCT arms	Retest A6MCT arms	A6MCT legs (rev.)	Z-scores A6MCT legs	Retest 6MCT legs	6MWT (m)	Z-scores 6MWT	MFM (%)	Brooke	Vignos	Walk/Run 10 m (s)	Climb 4 stairs (s)	Rise from floor (s)	Rise from chair (s)
1	F	CM, de novo AC7A1 mutation	10	10	32	139	16.6	4	4	Yes	520	-1.70	470	490	-4.24	624	212	-7.4	64	1	3	10.6*	16.8†	Impossible	1.2†
2	F	MD	8	5-6	25	126	15.8	5	5	No	498	-1.20	496	528	-3.09	536	425	-2.9	92	1	1	3.8	3.1	7.3	0.6
3	M	MD	15	9-11	58	172	19.6	5	5	No	472	-4.29	579	528	-4.33	585	503	-3.2	96	1	2	4.0*	1.9†	3.2	1.0
4	M	CM, fiber type disproportion	10	10	26	136	14.2	4	5	Yes	655	-1.26	611	478	-4.95	643	485	-2.6	97	1	1	3.1	1.5	2.3	1.2
5	M	SMA type 3	6	6	17	112	13.1	4	4	Yes	615	-0.51		535	-3.47		361	-2.6	92	1	2	5.1*	2.9†	1.9†	0.8
6	F	SMA type 3	9	9	21	125	13.4	4	4	Yes	583	-0.41		323	-6.71		NF		75	1	3	15.2*	9.6†	22.1†	1.4†
7	F	MD; an expansion > 150 CTG repeats in the DMPK gene	9	9	30	142	14.9	5	5	No	548	-1.76	616	688	-1.26	737	509	-2.0	99	1	1	3.3	1.5	1.8	0.9
8	M	MD; an expansion > 150 CTG repeats in the DMPK gene	7	7	24	123	15.9	5	5	No	604	-0.88	NF	603	-2.6†	NF	489	-1.2	100	1	1	2.9	1.3	1.3	0.9
9	M	CM, type I dominance	6	6	20	118	14.4	4	4	Yes	381	-2.6†	335	441	-4.47	416	NF		81	1	2	8.8*	5.3†	6.4†	0.8
10	F	CM, fiber type disproportion; TPM3 mutation	13	13	32	150	14.2	5	5	Yes	366	-4.30		424	-6.58		444	-4.0	97	1	1	3.7*	1.9	2.1	0.9
Median	-	-	9	-	26	130	14.6	-	-	-	534	Mean -1.89	538	509	Mean -4.17	605	434	Mean -3.2	94			3.9	2.4	2.3	0.9
(IQR)			(6.8-10.8)		(21-32)	(122-144)	(14.0-16.1)				(449-607)		(436-612)	(473-552)		(506-667)	(246-492)		(80-97)			(2.9-15.2)	(1.5-6.4)	(1.8-6.9)	(0.8-1.2)

\* = Abnormal gait pattern; † = Using hands. 6MWT = 6-minute walk test. A6MCT = Assisted 6-minute cycling test. BMI = Body Mass Index. cm = centimetres; CM = Congenital Myopathy; kg = kilogram; m = meters; MD = Myotonic Dystrophy; MFM = Motor Function Measure; MFC = Medical Research Council; NF = not feasible; Rev = revolutions; sec = seconds; SMA = Spinal Muscular Atrophy

Seventeen complication forms from eight patients with a OXPHOS dysfunction in muscle were returned to hospital. Only patient 1 and 2 reported no complaints after the test, though patient 1 reported mild tiredness after one out of three tests. The median scores for tiredness (IQR 1 - 2.5; range 0 - 3), muscle pain (IQR 0.5 - 3; range 0 - 3) and concentration problems (IQR 0 - 2.8; range 0 - 3) the day after the test was 2 (a bit more than normally). The median scores for muscle weakness (IQR 0 - 2; range 0 - 3), balance (IQR 0 - 1; range 0 - 2), coordination problems (IQR 0 - 1; range 0 - 2) and behavioural problems (IQR 0 - 2.5; range 0 - 3) were 1 (barely more than normally). Only three patients reported a 3 (a lot more than normally) tiredness, muscle weakness, muscle pain or behavioural problems, though only one patient reported a 3 consistently.

## Test results

### *Neuromuscular disorders*

Table 1A shows the test results for the patients with neuromuscular disorders. The median number of rev. cycled during the first test was 534 (IQR 449 - 607 rev.; range 366 - 655 rev.) with the arms and 509 (IQR 437 - 552 rev.; range 323 - 688 rev.) with the legs. The median walking distance with the 6MWT was 434 m (IQR 246 - 492 m; range 99 - 509 m).

### *OXPHOS dysfunction*

Table 1B shows the test results for the patients with OXPHOS dysfunction. The median number of rev. achieved during the A6MCT at the first visit was 600 rev. (IQR 542 - 748 rev.; range 499 - 823 rev.) with the arms, and 665 rev. (IQR 599 - 795 rev.; range 511 - 795 rev.) with the legs. The median walking distance in 6 minutes was 434 m (IQR 3 - 586 m; range 312 - 625 m). The median time for standing up from a chair and for climbing the stairs were not reliable because the timing mainly reflected the reaction time of the researcher.

## Validity

Although the limited number of patients included in this exploratory study will not allow us to draw firm conclusions, we aimed to get an impression of validity by testing the concept of submaximal exercise capacity and by correlating with motor capacities

### *Neuromuscular disorders*

The median percentage of the maximal heart rate (aim 70%) during the A6MCT<sub>arms</sub>, A6MCT<sub>legs</sub> and the 6MWT were 62%, 62% and 71% respectively. The median OMNI-scores during the 6<sup>th</sup> minute (aim 6) were 10 (IQR 3 - 10) for the arms and 3 (IQR 1 - 9) for the body during the A6MCT<sub>arms</sub> and 6 (IQR 2 - 10) for the legs and 3 (IQR 1 - 9) for the body during the A6MCT<sub>legs</sub>. The validity of this OMNI scale is, however, uncertain because many patients had difficulty indicating their tiredness and some OMNI scores fluctuated at the end of the test. The A6MCT<sub>legs</sub> did not correlate significantly to the



6MWT ( $p = 0.96$ ) or any of the sub domains of the MFM or timed tests ( $p = 0.13 - 0.70$ ;  $n = 8$ ). The A6MCT<sub>arms</sub> did not correlate statistically significantly with the 6MWT ( $p = 0.04$ ), any of the sub domains of the MFM or timed function tests ( $p = 0.42 - 0.88$ ). The slopes of the A6MCTs showed a stable number of rev. per minute (Figure 3A and 3B). In two patients with SMA type 3 we found near to normal results for the A6MCT<sub>arms</sub>, but not for the A6MCT<sub>legs</sub> which fits the disease course and gives an indication that both A6MCT are able to measure disease severity in upper and lower limb.

#### *OXPHOS dysfunction*

We found a median % of maximal heart rate (aim 70%) of 62% in both A6MCTs and 77% during the 6MWT. The OMNI score showed poor face validity (e.g. rating a 10 after 1 minute despite not looking tired) in most patients and is therefore not reported. The slopes of the A6MCTs showed a stable number of rev. per minute (Figure 3D and 3E). The A6MCT<sub>legs</sub> correlated strongly with the 6MWT ( $r = 0.83$ ;  $p = 0.042$ ), but not with the total and the individuals domain scores of the MFM ( $p = 0.05 - 0.63$ ) or the timed function tests ( $p = 0.20 - 0.51$ ). No significant correlations were found between the A6MCT<sub>arms</sub> and the 6MWT ( $p = 0.71$ ), any of the domains of the MFM ( $p = 0.18 - 0.40$ ), or timed function tests ( $p = 0.60 - 0.80$ ).

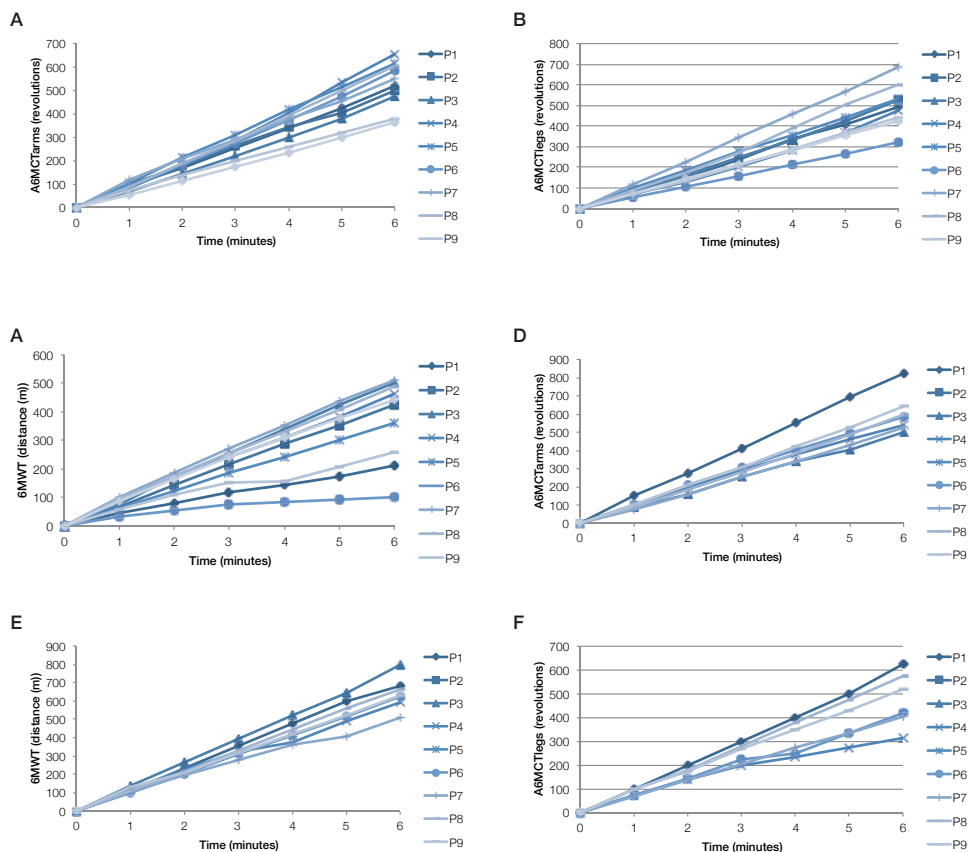
### Test-retest reliability

#### *Neuromuscular disorders*

Seven patients with neuromuscular disorders agreed to participate in the retest. The median time between test and retest was 14 days (IQR 13 - 20; range 12 - 21 days). The results of the retest of patient 8, in which the test was not feasible, were excluded from the analyses. The mean absolute differences between the scores of the test and the retest of the A6MCT<sub>arms</sub> and the A6MCT<sub>legs</sub> were 53 (min-max: -50 - +107) and 73 (min-max: -25 - + 165) rev., respectively. Five out of six children had a higher number of rev. during the retest for the legs, compared to two out of six for the arms. The ICC of the A6MCT was 0.77 (95% CI 0.05-0.97) for arms and 0.72 (95% CI -0.07 - 0.96) for the legs.

#### *OXPHOS dysfunction*

Five patients with OXPHOS dysfunction agreed to perform a retest. The median time between the test and the retest was 22 days (IQR 11 - 29 days; range 7 - 30 days). The mean absolute differences between the test and retest for the A6MCT<sub>arms</sub> and A6MCT<sub>legs</sub> were 57 (range 18 - 135) and 61 (range 6 - 105), respectively. The ICC for the A6MCT<sub>arms</sub> was 0.73 (95% CI -0.21 - 0.97). The ICC for the A6MCT<sub>legs</sub> was 0.91 (95% CI 0.37 - 0.99). The distance at the retest was higher compared to the initial test in the A6MCT<sub>arms</sub>, the A6MCT<sub>legs</sub> and the 6MWT for all children (average 34 and 110 rev. and 45 m, respectively).



**Figure 3.**

**Results of the 6-minute walking and assisted 6-minute cycling tests with arms and legs in patients with neuromuscular disorders and patients with OXPHOS dysfunction.** Number of revolutions for the arms (A) and legs (B) assessed using the A6MCT, and distance (meters) assessed using the 6MWT (C) per minute per patient for neuromuscular disorders. Number of revolutions for the arms (D) and legs (E) assessed using the A6MCT, and distance (meters) assessed using the 6MWT (F) per minute per patient with OXPHOS dysfunction. A6MCT = assisted 6-minute cycling test; 6MWT = 6-minute walk test.

## Discussion

In this study, we aimed to explore the feasibility, construct validity and test-retest reproducibility of the A6MCT for arms and legs in children with a neuromuscular disorder other than DMD, or OXPHOS dysfunction. The A6MCT seems to be a feasible test to measure submaximal exercise capacity in the majority of children in both ambulatory patient groups without moderate to severe cognitive impairment, behavioural problems or attention deficit.

In contrast to the 6MWT, which was feasible in only 70% of the children with a neuromuscular disorder and in 67% of the children with OXPHOS dysfunction, the A6MCT was feasible in 90% of the children with a neuromuscular disorder and in 78% of the patients with OXPHOS dysfunction. As expected, the A6MCT was not feasible in children with behavioural problems or severe developmental delay. Face validity was limited in two patients with OXPHOS dysfunction, with coordination problems and near-normal muscle strength, where the bicycle could not be stabilised during cycling with the legs, and it patients were not able to reach their submaximal exercise capacity (heart rate 56 and 61% of maximal heart rate). Good construct validity of the A6MCT was indicated by the moderate correlation between the A6MCT<sub>legs</sub> and the 6MWT in patients with OXPHOS dysfunction, but not in patients with neuromuscular disease. We could not confirm the correlation between the MFM and A6MCT<sub>legs</sub> found in patients with DMD.<sup>5</sup> Test-retest reliability of both A6MCTs was acceptable in both groups, which is in concordance with the data found by Jansen *et al.*,<sup>5</sup> and comparable with reliability results of the 6MWT in children.<sup>16,17</sup>

The submaximal character of the A6MCT was confirmed by a median percentage of maximal heart rate of 62% in both patients groups. The median heart rate was slightly but significantly lower during both A6MCTs compared to the 6MWT (71% for Neuromuscular Disorders and 77% for patients with OXPHOS dysfunction). However, since the maximum heart rate of children with neuromuscular disorder or with OXPHOS dysfunction during submaximal exercise has never been studied, these numbers should be interpreted cautiously. The constant slope of the number of rev. per minute curves of both A6MCT in all patients compared to the 6MWT curve, which showed deflecting curves in some patients, also suggests the ability of the A6MCTs to measure submaximal exercise capacity.

Although the A6MCT was developed for patients who are expected to lose their walking ability and for children who already have become wheelchair-dependent, we included less severely affected children to allow validation based on the results of the 6MWT. Especially the patients with OXPHOS dysfunction had near normal motor abilities, illustrated by the (near) normal MFM scores in all patients. The combination of the near

normal MFM and A6MCT results and the low heart rate during the A6MCTs suggests that the A6MCT was not strenuous enough for some of these ambulatory children. One should however keep in mind that the inclusion of ambulatory children with near normal motor abilities limits the external validity to non-ambulatory children or to children who are about to lose their ability to walk.

Since endurance and fatigability are important determinants of the patients' abilities to perform activities of daily life<sup>4</sup> and children with OXPHOS dysfunction and their parents rate fatigue and lack of energy as one of the most burdensome symptoms,<sup>18</sup> it seems reasonable to use outcome measure for endurance in future clinical trials for children with neuromuscular disorder and OXPHOS dysfunction. Since the most commonly used submaximal endurance test in children, the 6MWT, is not feasible in non-ambulatory children or children who are about to lose their ability to walk, the A6MCT seems a good alternative for these patients and should be studied in more detail.

The major weakness of this exploratory study includes the small number of children with heterogeneity clinical background not belonging to the target population for which the A6MCTs were developed. Future studies should therefore focus on validating the A6MCT tests in a larger group of patients within the targeted population (*i.e.* children who are about to lose their ability to walk or are already non-ambulatory). Since the 6MWT is not feasible in non-ambulatory patients, we suggest to use other functional tests to confirm construct validity. Since responsiveness was not tested in this study, we also suggest longitudinal studies in patients with progressive neuromuscular or OXPHOS dysfunction. For future studies in patients with OXPHOS dysfunction, we suggest to include patients with mitochondrial myopathies and below-normal exercise capacity on the 6MWT; alternatively the A6MCT can be used when the device is better fixed or more resistance is used. For the latter population, the possibility of a learning effect in the A6MCTs should also be evaluated.

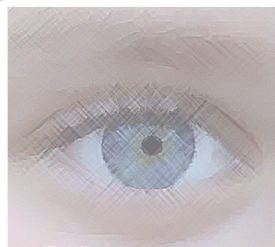
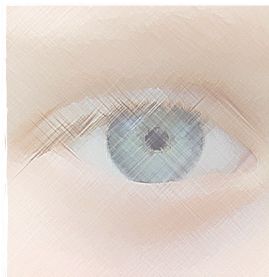
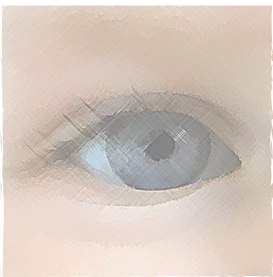
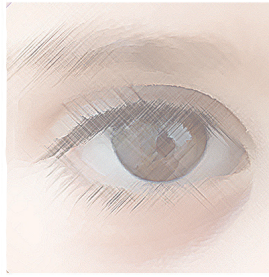
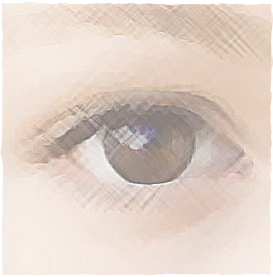
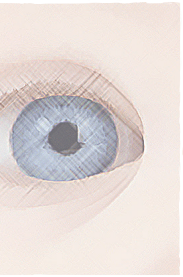
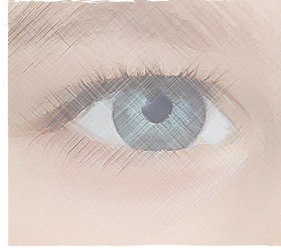
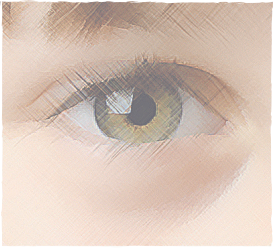
In conclusion, the A6MCTs seem a promising test to measure submaximal endurance and fatigability in children with a neuromuscular disorders and children with OXPHOS dysfunction. In children who are able to walk and are not expected to lose their ability to walk during the study, we still recommend the 6MWT.

## References

1. Laboratories ATSCoPSfCPF. ATS statement: guidelines for the six-minute walk test. *American journal of respiratory and critical care medicine* 2002;166:111-7.
2. Sartor F, Vernillo G, de Morree HM, *et al.* Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. *Sports medicine* 2013;43:865-73.
3. Bar-Or O. Role of exercise in the assessment and management of neuromuscular disease in children. *Medicine and science in sports and exercise* 1996;28:421-7.
4. Mazzone E, Martinelli D, Berardinelli A, *et al.* North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscul Disord* 2010;20:712-6.
5. Jansen M, De Jong M, Coes HM, Eggermont F, Van Alfen N, De Groot IJ. The assisted 6-minute cycling test to assess endurance in children with a neuromuscular disorder. *Muscle & nerve* 2012;46:520-30.
6. Brooke MH, Griggs RC, Mendell JR, Fenichel GM, Shumate JB, Pellegrino RJ. Clinical trial in Duchenne dystrophy. I. The design of the protocol. *Muscle & nerve* 1981;4:186-97.
7. Jansen M, de Groot IJ, van Alfen N, Geurts A. Physical training in boys with Duchenne Muscular Dystrophy: the protocol of the No Use is Disuse study. *BMC pediatrics* 2010;10:55.
8. McDonald CM, Henricson EK, Han JJ, *et al.* The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle & nerve* 2010;41:500-10.
9. Robertson RJ, Goss FL, Boer NF, *et al.* Children's OMNI scale of perceived exertion: mixed gender and race validation. *Medicine and science in sports and exercise* 2000;32:452-8.
10. Pfeiffer KA, Pivarnik JM, Womack CJ, Reeves MJ, Malina RM. Reliability and validity of the Borg and OMNI rating of perceived exertion scales in adolescent girls. *Medicine and science in sports and exercise* 2002;34:2057-61.
11. Angelini C, Semplicini C, Ravaglia S, *et al.* New motor outcome function measures in evaluation of late-onset Pompe disease before and after enzyme replacement therapy. *Muscle & nerve* 2012;45:831-4.
12. Florence JM, Pandya S, King WM, *et al.* Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. *Phys Ther* 1984;64:41-5.
13. Berard C, Payan C, Fermanian J, Girardot F. [A motor function measurement scale for neuromuscular diseases - description and validation study]. *Rev Neurol (Paris)* 2006;162:485-93.
14. Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord* 2005;15:463-70.
15. Munro BH. Correlation. In: Munro BHe, ed. *Statistical methods for health care research*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2005:239-58.
16. McDonald CM, Henricson EK, Abresch RT, *et al.* The 6-minute walk test and other clinical endpoints in duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicentre study. *Muscle & nerve* 2013;48:357-68.
17. Kempen JC, Harlaar J, van der Kooi AJ, *et al.* Reliability of the walking energy cost test and the six-minute walk test in boys with Duchenne muscular dystrophy. *Neuromuscul Disord* 2014;24:216-21.

18. Koene S, Wortmann SB, de Vries MC, *et al.* Developing outcome measures for pediatric mitochondrial disorders: which complaints and limitations are most burdensome to patients and their parents? *Mitochondrion* 2013;13:15-24.
19. Geiger R, Strasak A, Trembl B, Gasser K, Kleinsasser A, Fischer V, Geiger H, Loeckinger A, Stein JI. Six-minute walk test in children and adolescents. *J Pediatr.* 2007 Apr;150(4):395-9, 399.e1-2.







# Chapter 9

Is 2D speckle tracking echocardiography useful for detecting and monitoring myocardial dysfunction in adult m.3243A>G carriers? – a retrospective pilot study



## Is 2D speckle tracking echocardiography useful for detecting and monitoring myocardial dysfunction in adult m.3243A>G carriers? – a retrospective pilot study

**Saskia Koene**<sup>1</sup>, Janneke Timmermans<sup>2</sup>, Gert Weijers<sup>3</sup>, Paul de Laat<sup>1</sup>, Chris L. de Korte<sup>3</sup>, Jan A.M. Smeitink<sup>1</sup>, Mirian C.H. Janssen<sup>1,4</sup>, Livia Kapusta<sup>5,6</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Department of Paediatrics, Radboudumc

<sup>2</sup>Department of Cardiology, Radboudumc

<sup>3</sup>Clinical Physics Laboratory, Department of Radiology, Radboudumc

<sup>4</sup>Department of Internal Medicine, Radboudumc

<sup>5</sup>Department of Paediatrics, Paediatric cardiology Unit, Tel-Aviv Sourasky Medical Centre, Tel Aviv, Israel

<sup>6</sup>Children's Heart Centre, Radboudumc, Amalia Children's Hospital, Nijmegen, The Netherlands

Submitted

Cardiomyopathy is a common complication of mitochondrial disorders, associated with increased mortality. Two dimensional speckle tracking echocardiography (2DSTE) can be used to quantify myocardial deformation. Here, we aimed to determine the usefulness of 2DSTE in detecting and monitoring subtle changes in myocardial dysfunction in carriers of the 3243A>G mutation in mitochondrial DNA.

In this retrospective pilot study, thirty symptomatic and asymptomatic carriers of the mitochondrial 3243A>G mutation of whom two subsequent echocardiograms were available were included. We measured longitudinal, circumferential and radial strain using 2DSTE. Results were compared to published reference values.

Speckle tracking was feasible in 90% of the patients for longitudinal strain. Circumferential and radial strain showed low face validity (low number of images with sufficient quality; suboptimal tracking) and were therefore rejected for further analysis. Global longitudinal strain showed good face validity and was abnormal in 56 – 70% (depending on reference values used;  $n = 27$ ) of the carriers. Reproducibility was good (mean difference of 0.83 for inter- and 0.40 for intra-rater reproducibility; ICC 0.78 and 0.89, respectively). The difference between the first and the second measurement exceeded the measurement variance in 39% of the cases ( $n = 23$ ; feasibility of follow-up 77%).

In conclusion, two-dimensional strain echocardiography seems a feasible method to detect and monitor subtle changes in longitudinal myocardial deformation in adult carriers of the mitochondrial 3243A>G mutation. Based on our data and the reported accuracy of global longitudinal strain in other studies, we suggest the use of global longitudinal strain in a prospective intervention study.

## Introduction

Cardiomyopathy is a common complication of mitochondrial disorders, with a prevalence up to 14% in adults<sup>1</sup> and 58% in children.<sup>2,3</sup> The presence of cardiomyopathy has been associated with increased mortality in both children and adults.<sup>2-5</sup> In carriers of the m.3243A>G mutation, the most common genetic cause of mitochondrial disease,<sup>6</sup> the prevalence of either symptomatic or asymptomatic cardiomyopathy ranges from 18 to 56%, depending on the population studied and the method used.<sup>5,7,8</sup> Cardiomyopathy associated with this mutation was characterised as a concentric left ventricle hypertrophy<sup>8,9</sup> with progressive dilation and decreased systolic function, developing over several years.<sup>10</sup> Some, but not all of these small studies, identified more severe disease and high heteroplasmy percentage as factors contributing to cardiac involvement.<sup>7,8,11</sup>

The m.3243A>G mutation is a mutation in mitochondrial DNA and therefore follows maternal inheritance.<sup>12</sup> Of importance for understanding this study are the following aspects: the mutation is present in a variable percentage of all mtDNA (heteroplasmy), virtually all organs may be affected in various different patterns, and this is not only dependent on the heteroplasmy percentage of the mutation.

The severity of the myocardial dysfunction is usually assessed by conventional echocardiography and tagged cardiac magnetic resonance imaging (cMRI). Previously, tagged cMRI showed abnormal myocardial deformation in 22 asymptomatic m.3243A>G carriers.<sup>13</sup> cMRI is used as a reference standard for myocardial deformation, but it is a time-consuming and expensive procedure which requires dedicated expertise. Therefore, this method is less feasible as a bed side modality in multi-centre trials. Conventional echocardiography, in contrast, is widely available, but has low sensitivity in detecting subtle and regional changes in myocardial function.<sup>14</sup> The off-line processing of conventional echocardiograms using 2 dimensional speckle tracking echocardiography (2DSTE) software enables more sensitive quantification of the global and regional myocardial deformation<sup>15</sup>. This software tracks myocardial ‘speckle’ patterns throughout the myocardial cycle, facilitating calculation of myocardial deformation or strain in three directions (longitudinal, radial, and circumferential). The technique is reliable<sup>16</sup> and accurate<sup>17,18</sup> and is able to detect subtle changes in myocardial function at an early stage, even before decrease in conventional echocardiographic parameters (e.g. ejection fraction and shortening fraction) are observed.<sup>19</sup> Recently, global longitudinal strain, known as the most reliable component of strain analysis,<sup>20</sup> was incorporated into the recommendations for multi-modality imaging evaluation and monitoring of cardiac (dys)function of adult patients during and after cancer therapy.<sup>21</sup> A 10% decrease in longitudinal strain is a significant outcome measure for chemotherapy induced cardiomyopathy.<sup>21</sup>

In this retrospective pilot study we aim to evaluate the usefulness and feasibility of 2DSTE in detecting and monitoring subtle changes in myocardial deformation in adults carriers of the m.3243A>G mutation with a wide spectrum of clinical disease severity, since 2DSTE might be a potential outcome measure to evaluate responsiveness to future therapy for this group of carriers.

## Methods

### Study population

All subjects were identified from our 'National inventory of patients with the m.3243A>G mutation' study, including both symptomatic and asymptomatic carriers.<sup>22</sup> Subjects with a detectable heteroplasmy percentage (detection limit  $\geq 5\%$ ) in either buccal mucosa cells, leukocytes, or urinary epithelial cells (UEC) are considered to be carriers of the mutation. As part of usual clinical care, both symptomatic and asymptomatic carriers undergo cardiac ultrasound (approximately two-yearly in asymptomatic individuals; more frequently if clinically indicated). Carriers of the m.3243A>G mutation of whom two subsequent echocardiograms were available were included in this study. Only part of these patients were included in our previously published study on biomarkers.<sup>23</sup>

### Clinical assessment protocol

The carriers' clinical assessment at the time of the echocardiogram included the carrier's medical history, current cardiac complaints and the use of medication, physical examination including blood pressure, height and weight and an electrocardiogram (ECG).

In the context of the 'National inventory of patients with the m.3243A>G mutation', patients and their maternal relatives undergo several investigations including the assessment of general mitochondrial disease severity using the Newcastle Mitochondrial Disease Adult Scale (NMDAS).<sup>24</sup> The NMDAS score closest in time to the echocardiography was reported. The NMDAS contains the following four domains: i) current function; ii) system specific involvement; iii) current clinical assessment; and iv) quality of life. In our analysis, we used domains i-iii to calculate disease severity. Severe mitochondrial disease was defined previously as an NMDAS score above 20.<sup>22</sup> We defined asymptomatic disease as an NMDAS = 0; mild mitochondrial disease as an NMDAS score of 1 through 5; and moderate mitochondrial disease as an NMDAS score of 6 through 20.<sup>25</sup> The presence and severity of diabetes mellitus (DM) and cardiovascular involvement was also obtained from the NMDAS.<sup>26</sup> Quality of life, with sub scores for mental and physical quality of life, was determined using a Dutch translation of the SF-12v2 and American reference values.<sup>27</sup> The quality of life can vary from 0 to 70 for both mental

and physical health, where 50 is the population's mean (standard deviation 10). Subjective change during follow-up was part of the general history taking of the cardiologist.

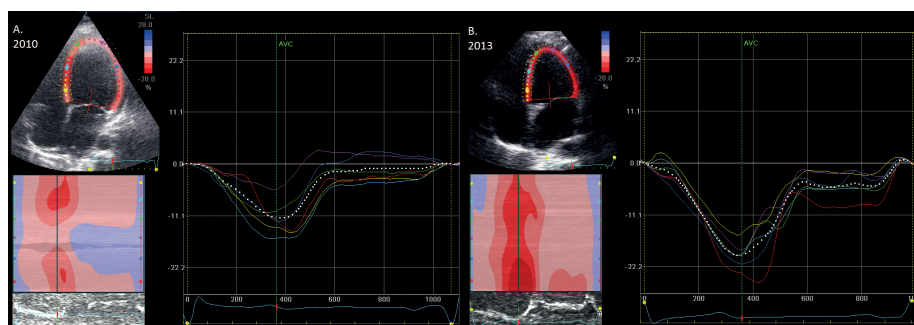
Laboratory investigations (+/- 6 months) were reported. Carriers were classified as having decreased creatinine clearance only, microalbuminuria only, both or neither. Microalbuminuria was defined as an albumin-to-creatinine-ratio of >2.0 g/mol for men and >2.5 g/mol for women. Decreased creatinine clearance was defined as a glomerular filtration rate <60 ml/min/1.73m<sup>2</sup>.

## Echocardiography

All carriers underwent a transthoracic 2D echocardiogram in supine and lateral position at rest. The echocardiogram was performed as part of our regular patient care (using a standardised echocardiographic protocol published previously<sup>28</sup>), by an experienced echocardiography technicians and supervised by experienced cardiologist. Images were obtained with a, M3S transducer using the Vivid 7 and M5S transducer using Vivid E9 echographic scanners (GE, Vingmed Ultrasound, Horten, Norway). Quantification of cardiac chamber size, left ventricular mass and systolic and diastolic left ventricular function were performed in accordance with the recommendations for chamber quantification by the American Society of Echocardiography's Guidelines and Standard Committee and the Chamber Quantification Writing Group<sup>29-31</sup> and were obtained from the 2D echocardiogram, as previously described.<sup>28</sup> In case the ejection fraction (EF) was not available, fractional shortening (FS) was used.

Strain analysis was done according to our previously published protocol,<sup>16</sup> by an experienced investigator (LK), using two-dimensional gray scale images taken in the parasternal apical 4-chamber view (4-CV) and at mid-cavity short-axis view (at the level of the papillary muscle; SAX-PM). The investigator was not aware of the clinical condition and medical treatment of the patients. A sector scan angle of 30-60 degrees was chosen and frame rates of 70 Hz or more were used.<sup>32</sup> Cine loops of preferably three cardiac cycles triggered by the R wave of the QRS complex were digitally saved. Offline analysis was performed using software for echocardiographic quantification (EchoPAC 6.1.0, GE Medical Systems, Horten, Norway). Timing of aortic valve closure and mitral valve opening was used to indicate end-systole and start of diastole respectively. Manual tracking of the endomyocardial borders was performed at the end-systolic frame. An automatic generation of the second epicardial tracing was created by the software, which also automatically divided the LV myocardium into six equal segments, which were named and localised according to the statement of the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association.<sup>33</sup> Quality of the speckle tracking was verified for each segment and adjusted when needed. Tracking was only accepted if visual inspection indicated adequate tracking over the full cardiac cycle. Preferably, three cardiac cycles were analysed for each segment and exported

as text files for further post-processing. As a next step the exported strain curves were post-processed, in a custom made software package using Matlab (r2013b), in order to estimate the average strain curve and to obtain the final strain parameters out of this data.<sup>34</sup> Figure 1 shows a composite figure with the region of interest (left side) and graphic depiction of longitudinal strain in 4-chamber long axis view and for radial strain in the mid-cavity short-axis view. The inter and intra-rater variability of our group was previously published.<sup>34</sup>



**Figure 1.**

**Two-dimensional strain measurement.** Two-dimensional strain measurement in longitudinal plane. At the left top, the region of interest is shown, at the right the graphic depiction of the transmural strain in a young female m.3243A>G carrier who was found to have severe cardiomyopathy in 2010. After diagnosis, she was started on medication (diuretics, ACE inhibitor and  $\beta$  blocker) and underwent an intensive heart failure rehabilitation programme. At the echocardiogram in 2013, she reported a highly significant increased in exercise tolerance. The coloured lines represent the measurements of regional deformation of the individual regions, the dotted line represents their average (global strain), analysed by GE EchoPac. GLS increased from -12.7 to -19.6.

To evaluate the conventional echocardiographic values obtained in our patient group, we used the reference values described by the American Society of Echocardiography.<sup>29,30</sup> Strain values are dimensionless and are expressed in percentage. The average values per segment of longitudinal, radial and circumferential end-systolic strain was determined.<sup>35</sup> End-systolic strain, the recommended measure for reporting myocardial deformation in a recent consensus paper,<sup>36</sup> is strain measured at end-systole, by time of aortic valve closure. Global longitudinal myocardial strain (GLS) was calculated by averaging the six segments using the apical 4-chamber view (4-CV). Global radial and circumferential myocardial strain (GRS and GCS, respectively) were calculated by averaging the six segments of the mid-cavity short-axis view (SAX-PM). When less than four segments were available for averaging, global strain was not calculated. In a recent consensus paper by Voigt *et al.* end-systolic strain (at the time of the aortic valve closure) was strongly recommended as “the measure” for reporting myocardial deformation.<sup>36</sup> Since the end-systolic strain value may differ from peak systolic strain,



which is a measurement of peak deformation at any point during systole, we only used published reference values for end-systolic strain. Age-matched reference values for end-systolic strain were obtained from Kocabay *et al.*<sup>20</sup> and from Kuznetsova *et al.*<sup>37</sup> Values were considered abnormal when deviating >2 standard deviations (SD) from mean (Kocabay) or < p5 or > p95 (Kuznetsova, SDs were not reported). Each carrier was compared individually to his/her age-matched reference values. Changes were reported as absolute numbers and changes of  $\geq 10\%$  were accepted as real changes, as this number exceeds the coefficient of variation reported in literature<sup>38</sup> recommended by Plana *et al.*

### Reproducibility studies

Ten images in which strain analysis was possible were randomly selected by a physician not involved in the strain measurements for reproducibility studies. To determine inter-observer reproducibility, longitudinal strain was analysed by another experienced rater (GW), blinded to previous results. Intra-rater reproducibility was determined by rating the same images, more than six months later. For both inter- and intra-observer reproducibility, absolute differences and the intraclass correlation coefficients were calculated.

### Follow-up study

Only carriers of whom two echocardiograms in which GLS analysis was feasible were available were included in this study. Subjective and objective changes in clinical status, changes in the use of medication and changes in conventional parameters were assessed during follow-up. Since all echocardiograms were performed as part of clinical care, the time between two examinations varies.

### Medical ethical approval

This study is part of the “National Inventory of Patients with the m.3243A>G mutation”, which was approved by the regional Medical Research Ethics Committee. In accordance with the Helsinki agreement, written informed consent was obtained from each participant.

### Statistical analysis

The absolute difference was calculated by subtracting the first measurement from the second measurement. All variables were assessed for (log)normality. To prevent non-real values for zero values of the NMDAS including its sub domains and symptom specific items, these values were increased by 1 prior to elog transformation. Variables with a (log)normal distribution, were compared using parametric tests, and the mean and 95%-confidence intervals (95%CI) are reported. Variables that deviated strongly from a (log)normal distribution were analysed by performing a non-parametric test and the median and interquartile ranges (IQRs) are reported. Inter- and intra-rater reliability

were calculated using intraclass correlation coefficients for absolute agreement (ICC). Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website (<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>).

All analyses were performed using IBM's SPSS statistics software packages, version 20.0.0.1.

## Results

### Carrier description

The study algorithm is presented in Figure 2. Thirty carriers of the m.3243A>G mutation of whom two subsequent echocardiograms were available for 2D strain analysis were included in this study (Table 1). Nineteen of these carriers were female; five of them were current smokers. Eighteen carriers had diabetes mellitus, six microalbuminuria, one decreased creatinine clearance and ten carriers had cardiovascular involvement according to the NMDAS.

In the 'National Inventory of Patients with the m.3243A>G mutation', eighty adult carriers had been included before 2012 (since the regular follow-up time is 2 years, we do not account for carriers included before 2012). For a variety of reasons (*e.g.*, only one echocardiogram available, or echocardiogram made using a wrong echocardiograph device), only data of 30 carriers (38%) from 20 families could be analysed (1-3 family members per family). These 30 carriers were not different compared to the total cohort with respect to their total NMDAS score ( $p = 0.35$ ), nor in their sub scores for diabetes mellitus or cardiovascular disease ( $p = 0.37$  and  $0.81$  respectively). BMI was significantly higher in the included carriers compared to the other carriers ( $p = 0.0062$ ).

**Table 1.**

**Carrier characteristics.** Clinical features of the included carriers compared to all adult carriers in the 'National inventory of carriers with the m.3243A>G mutation'. The presence of Diabetes Mellitus was obtained from the NMDAS scale (score on Diabetes Mellitus item  $\geq 3$ ); the presence of cardiovascular involvement was obtained from the NMDAS scale (score on Cardiomyopathy  $\geq 1$ ).  $p$ -values for the difference at baseline between this cohort and all adult carriers included before 2012 were calculated.

		<i>n</i> =	mean (range)	total number of patients	Difference from total cohort of carriers ( <i>n</i> = 80) <i>p</i> =
Gender	Female	19		30	
Age			45.2 (16.7 - 64.5)	30	0.90
BMI	(kg/m <sup>2</sup> )		23.8 (17.3 - 34.0)	30	0.006
Alcohol	Use	6		27	
	Abuse	1		27	
Smoking	Current	5		27	
	Stopped	5		27	
Renal abnormalities	Micro-albuminuria	6		30	
	Decreased creatinine clearance	1		30	
Diabetes Mellitus (NMDAS)	Yes	18		30	0.37
HbA1c	(mmol/mol)		6.7 (5 - 9.7)	28	
Cardiovascular involvement (NMDAS)	Yes	10		30	0.84
Systolic blood pressure			127 (95 - 167)	23	
	<120 mmHg	5			
	120 - 140 mmHg	13			
	> 140 mmHg	5			
Diastolic blood pressure	(mmHg)		79 (58 - 98)	23	
Total cholesterol	(mmol/l)		4.8° (3.5 - 9.2)	23	
HDL cholesterol	(mmol/l)		1.1 (0.6 - 1.9)	21	
LDL cholesterol	(mmol/l)		2.9 (0.9 - 4.8)	21	
Heteroplasmy in leucocytes	(%)		28 (3 - 73)	30	
Heteroplasmy in UEC	(%)		52 (7 - 96)	30	
NMDAS			18.6 (0 - 49)	30	0.35
Domain 1			7.5 (0 - 21)	30	0.11
Domain 2			7.3 (0 - 18)	30	0.20
Domain 3			2.6* (0 - 13)	30	0.96
QoL Mental			46 (25 - 59)	28	0.34
QoL Physical			40 (21 - 58)	28	0.09
Medication <sup>o</sup>	β blocker	6		30	
	Calcium channel blocker	1		30	
	ACE inhibitor	6		30	
	Angiotensin II blocker	2		30	
	Insulin	5		30	

BMI = body mass index; Domain 1 = current function; Domain 2 = system specific involvement; Domain 3 = current clinical assessment; IQR = interquartile range; *n* = number of carriers of which data were available at that specific time point; NMDAS = Newcastle Mitochondrial Disease Adult Scale; QoL = Quality of Life; UEC = urinary epithelial cells. \*lognormal distribution; ° median is given instead of mean; <sup>o</sup> not mutually exclusive

## Conventional echocardiography at baseline

Table 2 presents the echocardiographic parameters of our patients. Using conventional echocardiography, seven carriers (23%) did not present with (sub-clinical) cardiac abnormalities. Hypertrophy was reported in 40% of the carriers, mild mitral insufficiency in 13% and mild aortic insufficiency in 27% of the carriers.

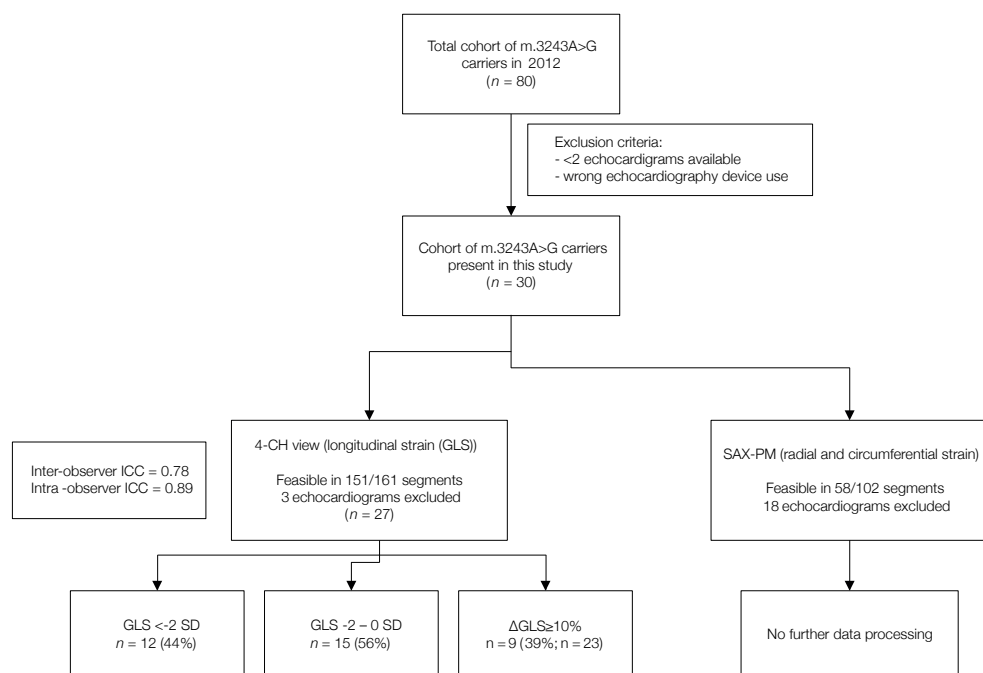
**Table 2.**

**Cardiac characteristics at baseline.** Conventional echocardiographic and myocardial deformation in m.3243A>G carriers compared to the reference population. Reference values were obtained from Nagueh *et al.* en Lang *et al.* for the conventional echocardiographic parameters and to Kocabay *et al.* and Kuznetsova *et al.* for strain data. High (>+2SD) and low (<-2SD) are based on age-matched reference values.

	median (spread)	High values (n (%))	Low values (n (%))	n =
FS (%)	36 (11 - 56)	2 (7%)	10 (33%)	30
EF (%)	64 (25 - 68)	-	3 (30%)	10
Interventricular septum thickness in diastole (cm)	0.9 (0.6 - 3.0)	12 (41%)	-	29
LV posterior wall thickness (cm)	0.9 (0.7 - 1.4)	13 (45%)	-	29
LV internal diameter in diastole (cm)	4.4 (1.9 - 6.0)	1 (3%)	6 (21%)	29
Mitral valve E/A Ratio	1.1 (0.7 - 4.0)	2 (7%)	1 (3%)	29
LV performance (Tei) index	0.4 (0.3 - 0.8)	11 (41%)	-	27
Isovolumic relaxation time (ms)	84 (50 - 120)	9 (32%)	2 (7%)	28
Pulmonary Vein S/D Ratio	1.4 (0.54 - 2.2)	5 (21%)	-	24
Left ventricular mass index (g/m <sup>2</sup> )	76 (23 - 140)	4 (17%)	2 (8%)	24
Global longitudinal strain compared to Kocabay <i>et al.</i>	-16.3 (-7.9 - -20.7)	-	15 (56%)	27
compared to Kuznetsova <i>et al.</i>		-	19 (70%)	27

E/A ratio = ratio between early (E) and late (A) filling velocity, measured at the mitral valve; LV = left ventricular, S/D ratio = ratio between the velocity of the flow in systole and diastole, □ not mutually exclusive, ° median and IQR are given instead of mean and 95%CI

The previously observed correlation between the left ventricular mass index and the NMDAS and heteroplasmy percentage<sup>9</sup> could not be confirmed in our cohort ( $r(\text{heteroplasmy UEC}) = -0.01$ ;  $p = 0.97$ ;  $r(\text{heteroplasmy leucocytes}) = -0.04$ ,  $p = 0.87$ ;  $r(\text{NMDAS}) = 0.01$ ,  $p = 0.97$ ;  $n = 24$ ).



**Figure 2.**  
**Study algorithm.**

4-CV = 4 chamber view; ΔGLS = change in global longitudinal end-systolic strain; GLS = global longitudinal end-systolic strain; ICC = intraclass correlation coefficient; SAX-PM = short axis view at level of papillary muscle

## Strain feasibility

For the analysis of GLS on the 4-CV images, three patients (10%) had to be excluded because no GLS analysis could be performed. Two instead of three cycles were suitable for analyses in 33% (e.g., bad quality of the images, only two cardiac cycles recorded). The assessment of GLS was feasible in 151 (94%) of the 161 segments available for analysis. During follow-up; two subsequent images could be analysed in 23 patients (feasibility 77%).

For the SAX-PM view, only 57% of the segments had sufficient image quality to perform radial and circumferential strain analyses and 18 patients (30%) had to be excluded because of low image quality. Because of the low number of high-quality images, the suboptimal tracking performance of some specific segments (mainly posterior and lateral)

even in the high-quality images, and the current debate on the reproducibility of radial strain," we chose not to further process these results and concentrate only on the longitudinal 2D strain.

### Longitudinal 2D strain at baseline

When comparing our results to the age-matched reference values of Kocabay *et al.*, 15 carriers (56%) had abnormal GLS (Table 3). When comparing to the age-matched reference values of Kuznetsova, 19 carriers (70%) had abnormal GLS. Of the seven carriers with normal cardiac function and diameters assessed by conventional echocardiography, two or four (depending on the reference values used) had abnormal GLS. In the ten carriers with low fractional shortening, 5 had low GLS (one GLS not available). In the nine carriers with a high diastolic left ventricular posterior wall thickness, seven had decreased GLS.

**Table 3.**  
**Global strain in carriers individually compared to age-matched controls.** Global longitudinal end-systolic strain in m.3243A>G compared to the age-matched reference population (*n* = 27). Marked in blue the low values (under -2 SD for age-matched reference values), marked in gray the values below the mean (between 0 and -2 SD). In green the high values (above +2 SD).

Sex	Age	Hetero- plasm Leuco- cytes (%)	Hetero- plasm UEC (%)	NMDAS	FS (%)	EF (%)	LVIDd (cm)	LVPWd (cm)	IVSd (cm)	MV E/A Ratio	GLS#	GLS*
M	29	51	98	49		25				1.08	-7.9	-7.9
F	30	42	ND	16	18		6	1	0.7	4	-12.7	-12.7
F	31	49	57	13	11		1.9	0.9	1.3	0.99	-12.8	-12.8
M	43	19	ND	7	20		3.9	1	1	0.8	-13.1	-13.1
F	47	31	61	43	17		4.4	0.9	0.8	1.26	-13.2	-13.2
F	58	17	41	21	37	63	3.8	1.3	1.8	1.49	-13.9	-13.9
F	38	23	30	21	44		4.4	1.2	1.4		-13.9	-13.9
M	51	16	78	16		51	4.4	1.3	1.4	0.91	-14.1	-14.1
F	57	25	85	15	49		3.5	1.1	1	0.75	-14.2	-14.2
F	49	17	45	38	33	67	3.8	1.4	1.3	0.86	-14.3	-14.3
M	39	30	54	11	43		5.6	0.9	1	1.44	-16.1	-16.1
F	28	27	67	8	27	66	4.8	0.9	0.9	0.71	-16.1	-16.1
F	60	6	30	26	40		4.6	1	1	1.01	-16.9	-16.9
M	54	6	77	15	30		4.5	0.9	1	1.16	-17.1	-17.1
F	26	34	86	0	30	66	4.4	0.9	0.8	2.22	-17.1	-17.1
F	56	10	27	26	24		4.5	0.9	0.9	0.73	-17.6	-17.6
F	46	19	44	21	15	48	4	0.9	0.6	0.89	-17.7	-17.7
M	17	19	58	0	37		5.3	0.8	0.8	1.97	-17.7	-17.7

M	40	17	84	7			4.5	0.9	0.8	1.62	-18.3	-18.3
M	61	16	79	20	37		4.7	1.3	1.2	1.09	-18.6	-18.6
M	25	28	64	0	36	63	5.2	0.8	3	3.04	-18.9	-18.9
F	53	8	38	24	39	67	4.7	0.8	0.6	1.1	-19.4	-19.4
F	65	7	52	39	43		4.8	0.7	0.9	0.9	-19.4	-19.4
F	39	27	79	17	21		4	0.9	0.7	0.93	-19.6	-19.6
F	57	12	86	25	37	68	4.7	0.9	0.8	1.4	-19.9	-19.9
F	33	14	32	10	56		4.5	0.8	0.8	1.26	-20	-20
F	48	17	52	25	39		3.8	1	0.9	0.95	-20.7	-20.7

E/A ratio = ratio between early (E) and late (A) filling velocity, measured at the mitral valve; F = female; FS = fractional shortening; GLS = global longitudinal end-systolic strain; GRS = global radial end-systolic strain; LVPWd = diastolic left ventricular posterior wall thickness; M = male; NMDAS = Newcastle Mitochondrial Disease Adult Scale; UEC = urinary epithelial cells; \* = Kocabay reference values (95%CI); # Kuznetsova reference values (90%CI)

No significant correlation between age and GLS was found ( $r = -0.21$ ;  $p = 0.30$ ). GLS was not significantly lower in carriers with DM ( $p = 0.68$ ). GLS was not significantly lower in carriers with a score 0 versus  $\geq 1$  on the cardiovascular involvement item in the NMDAS ( $p = 0.10$ ). There was no correlation between IVSd and GLS ( $p = 0.35$ ;  $p = 0.08$ ). Carriers with renal abnormalities (including microalbuminuria) did not have significantly lower GLS compared to carriers without renal abnormalities ( $p = 0.65$ ).

GLS correlated significantly to the heteroplasmy percentage in UEC, but not to heteroplasmy percentage in leucocytes ( $r = 0.45$ ;  $p = 0.05$  and  $r = -0.17$ ;  $p = 0.36$ , respectively). No significant correlation between the NMDAS score and GLS was found ( $r = 0.29$ ,  $p = 0.17$ ). The score in the first, subjective domain of the NMDAS did not correlate significantly to GLS ( $r = 0.18$ ,  $p = 0.41$ ). There was no significant difference in the GLS between carriers with asymptomatic, mild, moderate and severe general mitochondrial disease severity ( $p = 0.65$ ). Physical quality of life did not correlate significantly to GLS ( $r = 0.18$ ,  $p = 0.38$ ). No significant correlation between mental quality of life and GLS was found ( $r = 0.15$ ,  $p = 0.49$ ).

### Inter- and intrarater reliability

The ICC of inter-rater reliability was 0.78, with a mean difference of 0.83 (95%CI 0.38 – 1.85). The ICC of intra-rater reliability was 0.89, with a mean difference of 0.40 (95%CI 0.11 – 1.42).

### Follow-up

Twenty-three patients were suitable for the analysis of changes in myocardial deformation during follow-up (feasibility 77%). Table 4 shows the changes in GLS, the subjective changes in exercise tolerance and changes in medication. The median time between the first and the second echocardiogram was 2.0 years (IQR 1.1 – 2.7 years; range 0.5 – 4.6

years). The disease severity (including the Cardiovascular and Diabetes Mellitus score and sub domains) as well as the quality of life didn't change significantly during follow-up ( $p = 0.11 - 0.92$ ).

Nine carriers (39%) had a change in GLS  $\geq 10\%$ . The change in GLS ranged from -7.1 to +6.9 (mean 0.33; 95%CI -5.5 - +2.3), where a higher (thus positive) value represents an improvement in longitudinal strain. Nine carriers (39%) had a change in GLS  $\geq 10\%$ .

Examples of the difference in responsiveness of myocardial strain include two cases of young carriers with m.3243A>G related cardiomyopathy which were documented in more detail: One young woman with a generally mild phenotype and newly discovered severe heart failure at the first echocardiogram recovered very well under pharmacotherapy and rehabilitation (GLS -12.7 to -19.6), while a male carrier with MELAS syndrome had a stable, severe heart failure (GLS -7.9 to -5.4). See Figure 1 for the first case.

**Table 4.**  
**Clinical and strain parameters during follow-up in m.3243A>G carriers. Description of the change during follow-up and the change in global end-systolic strain in longitudinal and radial direction at baseline and during follow-up in m.3243A>G carriers ( $n = 23$ ).** Time between echocardiograms, subjective change in exercise tolerance and changes in medication are also depicted. Blue marking indicates a  $\geq 10\%$  increase in myocardial strain; green marks a  $\geq 10\%$  decrease in myocardial strain during follow-up.

Time between echocardiograms (years)	Subjective change	Beta blocker	Calcium channel blocker	ACE inhibitor	Angiotensin 2 receptor blocker	Thiazide diuretic	alpha blocker	Loop diuretic	Potassium-sparing diuretic	Insulin	GLS-1	GLS-2	$\Delta$ GLS
3.0	↓	x		x							-16.1	-9.1	7.1
2.9		x	x	x		x	x			x	-16.9	-13.9	3.1
		x		x		x				x			
2.9	↓									x	-19.9	-17.2	2.7
1.6		x		x				x	x		-7.9	-5.4	2.5
		x		x				x	x	x			
1.1	-										-17.6	-15.9	1.7
2.9		x	x								-19.4	-17.8	1.6
2.7				x						x	-20.7	-19.3	1.3
				x						x			
1.0	=	x					x				-17.1	-16.3	0.8
1.4	↓	x			x					x	-13.9	-13.5	0.4
		x								x			



## Discussion

276

A previous study showed that myocardial deformation, measured by cMRI, in m.3243A>G mutation carriers without known clinical cardiac involvement showed abnormal longitudinal shortening, whereas radial and circumferential strain were comparable to matched healthy controls.<sup>9</sup> This is in line with studies in other populations where GLS by echocardiography is reported to be most sensitive strain parameter to assess systolic dysfunction.<sup>19,40</sup> Global longitudinal strain is therefore accepted as a major outcome parameter of *e.g.* chemotherapy related cardiac dysfunction.<sup>21</sup> In general, GLS is reported to be relatively easy to measure and more consistent and reproducible compared to GCS.<sup>20</sup> The hypothesis is that GLS is affected at first because the longitudinally orientated subendocardial fibres are most susceptible to injury.<sup>19</sup> In our present study, we found no correlation between the longitudinal myocardial deformation and the clinical parameters or clinical scoring such as the NMDAS. Nevertheless, one should keep in mind that mild general mitochondrial disease may be associated with severe cardiomyopathy and therefore all carriers of the m.3243A>G mutation should be screened for any sign of (subclinical) cardiomyopathy.

In another follow-up study of m.3243A>G carriers, cMRI was used to monitor cardiac adaptations and safety of endurance training. No difference was found in GCS during follow-up; however, GLS was not assessed.<sup>41</sup> Other studies report good responsiveness of radial (and in lesser extent of longitudinal) strain,<sup>42</sup> and improvement of longitudinal strain under treatment<sup>43</sup> while others report no changes in myocardial strain during the follow-up of patients with progressive (non-mitochondrial) diseases.<sup>44</sup> In the present pilot study, we found changes in GLS exceeding the inter- and intra-observer variability in 39% of our cohort. Since we were not able to measure the disease progression with other methods rather than 2DSTE itself, one can not rule out that these changes in GLS could still partly represent the influence of other factors, *e.g.* changes in medication, or treatment of the most common cause of cardiac hypertrophy: hypertension. The influence of covariates of diminished myocardial deformation, including the decreased myocardial deformation associated with physiological aging<sup>45</sup> and with the presence of diabetes mellitus,<sup>46</sup> was not significant in our cohort. This is probably caused by the high number of young subjects with severe m.3243A>G associated cardiomyopathy included in our cohort and the lack of correlation between the presence of m.3243A>G associated cardiomyopathy and m.3243A>G associated symptoms in general and m.3243A>G associated diabetes mellitus, specifically. Although we found a moderate correlation between GLS and the heteroplasmy percentage measured in urinary epithelial cells, the heteroplasmy percentages in these tissues do not always represent the heteroplasmy percentage measured in cardiac muscle.<sup>4</sup>

Tagged cMRI is considered the reference standard for measuring myocardial deformation,<sup>47</sup> yet the lack of availability and expertise in many centres hampers its use in future clinical trials. Although 2DSTE is dependent on operator experience, machine settings

and the acoustic window for transducer placement of the patient,<sup>47</sup> the technique is recently often used in multi-centre clinical trials. Pitfalls of the 2DSTE method itself include its dependence on high quality images, frame rate and pre- and afterload, the high labour-intensity, and the limited standardisation between vendors. Weaknesses of this study include the data gathering as part of clinical care, the relatively small number of carriers of whom two echocardiograms were available for further 2DSTE, and the lack of standardised follow-up protocol to allow firm conclusions about the responsiveness of 2DSTE. Despite the fact that the 2D images were obtained as part of clinical care, the proportion of high quality images on which analysis was feasible was comparable to the literature.<sup>48</sup>

More and more trials are performed in patients with mitochondrial disease. There is an urgent need for robust outcome measures.<sup>49</sup> The wide availability, the non-invasive nature and independence of voluntary effort of 2DSTE makes it promising as an objective and quantitative end point for clinical trials. The data for this retrospective pilot study were collected as part of routine care (using a standardised imaging protocol) and only 77% of the patients were suitable for the assessment of longitudinal myocardial strain using 2DSTE. Since these numbers are probably higher in highly standardised prospective studies, we recommend testing the feasibility and responsiveness of GLS in more detail in a prospective (intervention) study. Myocardial deformation reflects not only the myocardial contractility, but probably also global forces such as pre- and afterload. For proper use in future clinical trials, we suggest standardisation of e.g. medications and fluid intake at the time of the echocardiography.

Given the high prevalence of cardiac abnormalities and the lack of clear clinical predictors for the presence of not-yet-symptomatic cardiomyopathy, screening for this condition might be of major importance. Whether early detection of myocardial dysfunction in these patients will be beneficial for the long term prognosis in this population, remains to be further clarified. Although myocardial deformation seems to be predictive of heart failure<sup>50</sup> and all-cause mortality,<sup>51</sup> the value of treating subtle sub clinical changes in myocardial deformation, detected using 2DSTE, should be studied prospectively in a larger cohort of adult m.3243A>G carriers.

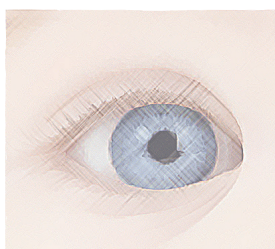
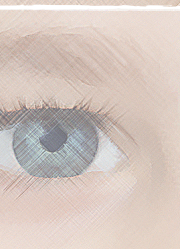
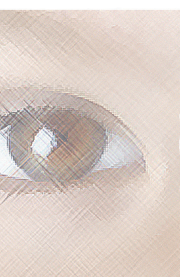
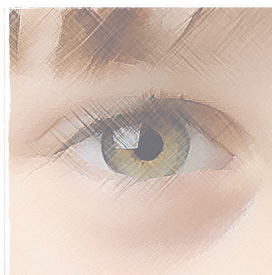
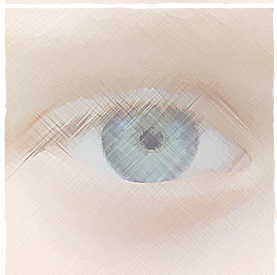
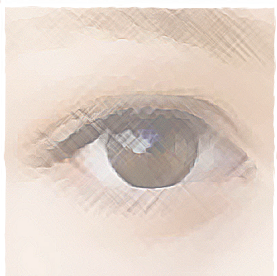
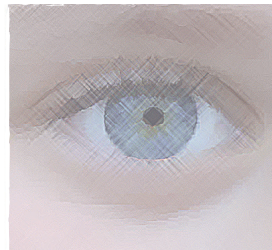
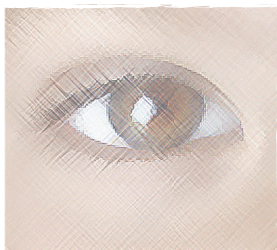
## References

1. Arpa J, Cruz-Martinez A, Campos Y, *et al.* Prevalence and progression of mitochondrial diseases: a study of 50 patients. *Muscle & nerve* 2003;28:690-5.
2. Scaglia F, Towbin JA, Craigen WJ, *et al.* Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics* 2004;114:925-31.
3. Holmgren D, Wahlander H, Eriksson BO, Oldfors A, Holme E, Tulinius M. Cardiomyopathy in children with mitochondrial disease; clinical course and cardiological findings. *Eur Heart J* 2003;24:280-8.
4. Majamaa-Voltti K, Turkka J, Kortelainen ML, Huikuri H, Majamaa K. Causes of death in pedigrees with the 3243A>G mutation in mitochondrial DNA. *Journal of neurology, neurosurgery, and psychiatry* 2008;79:209-11.
5. Malfatti E, Laforet P, Jardel C, *et al.* High risk of severe cardiac adverse events in patients with mitochondrial m.3243A>G mutation. *Neurology* 2013;80:100-5.
6. Greaves LC, Taylor RW. Mitochondrial DNA mutations in human disease. *IUBMB Life* 2006;58:143-51.
7. Vydut TC, de Coo RF, Soliman OI, *et al.* Cardiac involvement in adults with m.3243A>G MELAS gene mutation. *Am J Cardiol* 2007;99:264-9.
8. Majamaa-Voltti K, Peuhkurinen K, Kortelainen ML, Hassinen IE, Majamaa K. Cardiac abnormalities in patients with mitochondrial DNA mutation 3243A>G. *BMC cardiovascular disorders* 2002;2:12.
9. Bates MG, Hollingsworth KG, Newman JH, *et al.* Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers. *European heart journal cardiovascular Imaging* 2013;14:650-8.
10. Okajima Y, Tanabe Y, Takayanagi M, Aotsuka H. A follow up study of myocardial involvement in patients with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS). *Heart* 1998;80:292-5.
11. Hollingsworth KG, Gorman GS, Trenell MI, *et al.* Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load. *Neuromuscul Disord* 2012;22:592-6.
12. Smeitink JA, Zeviani M, Turnbull DM, Jacobs HT. Mitochondrial medicine: a metabolic perspective on the pathology of oxidative phosphorylation disorders. *Cell metabolism* 2006;3:9-13.
13. Bates MG, Hollingsworth KG, Newman JH, *et al.* Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers: defining cardiac adaptations and safety of endurance training in patients with m.3243A>G-related mitochondrial disease. *Eur Heart J Cardiovasc Imaging* 2013;14:650-8.
14. Aurigemma GP, Silver KH, Priest MA, Gaasch WH. Geometric changes allow normal ejection fraction despite depressed myocardial shortening in hypertensive left ventricular hypertrophy. *J Am Coll Cardiol* 1995;26:195-202.
15. Artis NJ, Oxenburgh DL, Williams G, Pepper CB, Tan LB. Two-dimensional strain imaging: a new echocardiographic advance with research and clinical applications. *Int J Cardiol* 2008;123:240-8.
16. Mavinkurve-Groothuis AM, Weijers G, Groot-Loonen J, *et al.* Interobserver, intraobserver and inpatient reliability scores of myocardial strain imaging with 2-d echocardiography in patients treated with anthracyclines. *Ultrasound Med Biol* 2009;35:697-704.

17. Amundsen BH, Helle-Valle T, Edvardsen T, *et al.* Noninvasive myocardial strain measurement by speckle tracking echocardiography: validation against sonomicrometry and tagged magnetic resonance imaging. *J Am Coll Cardiol* 2006;47:789-93.
18. Choi EY, Rosen BD, Fernandes VR, *et al.* Prognostic value of myocardial circumferential strain for incident heart failure and cardiovascular events in asymptomatic individuals: the Multi-Ethnic Study of Atherosclerosis. *Eur Heart J* 2013;34:2354-61.
19. Nesbitt GC, Mankad S, Oh JK. Strain imaging in echocardiography: methods and clinical applications. *Int J Cardiovasc Imaging* 2009;25:9-22.
20. Kocabay G, Muraru D, Peluso D, *et al.* Normal left ventricular mechanics by two-dimensional speckle-tracking echocardiography. Reference values in healthy adults. *Revista espanola de cardiologia* 2014;67:651-8.
21. Plana JC, Galderisi M, Barac A, *et al.* Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* 2014;15:1063-93.
22. de Laat P, Koene S, van den Heuvel LP, Rodenburg RJ, Janssen MC, Smeitink JA. Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A>G mutation. *Journal of inherited metabolic disease* 2012.
23. Koene S, de Laat P, van Tienoven DH, *et al.* Serum GDF15 Levels Correlate to Mitochondrial Disease Severity and Myocardial Strain, but Not to Disease Progression in Adult m.3243A>G Carriers. *JIMD reports* 2015;24:69-81.
24. Phoenix C, Schaefer AM, Elson JL, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord* 2006;16:814-20.
25. Koene S, de Laat P, van Tienoven DH, *et al.* Serum FGF21 levels in adult m.3243A>G carriers: Clinical implications. *Neurology* 2014;6:578.
26. Schaefer AM, Phoenix C, Elson JL, McFarland R, Chinnery PF, Turnbull DM. Mitochondrial disease in adults: a scale to monitor progression and treatment. *Neurology* 2006;66:1932-4.
27. QualityMetric: SF-12v2 Health Survey. April 11, 2006, at <http://www.qualitymetric.com/products/surveys/SF12v2.shtml> )
28. Bulten BF, Mavinkurve-Groothuis AM, de Geus-Oei LF, *et al.* Early myocardial deformation abnormalities in breast cancer survivors. *Breast cancer research and treatment* 2014;146:127-35.
29. Lang RM, Bierig M, Devereux RB, *et al.* Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography* 2005;18:1440-63.
30. Nagueh SF, Appleton CP, Gillebert TC, *et al.* Recommendations for the evaluation of left ventricular diastolic function by echocardiography. *European journal of echocardiography : the journal of the Working Group on Echocardiography of the European Society of Cardiology* 2009;10:165-93.
31. Lang RM, Bierig M, Devereux RB, *et al.* Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *J Am Soc Echocardiogr* 2005;18:1440-63.

32. Leitman M, Lysyansky P, Sidenko S, *et al.* Two-dimensional strain-a novel software for real-time quantitative echocardiographic assessment of myocardial function. *J Am Soc Echocardiogr* 2004;17:1021-9.
33. Cerqueira MD, Weissman NJ, Dilsizian V, *et al.* Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation* 2002;105:539-42.
34. Mavinkurve-Groothuis AM, Weijers G, Groot-Loonen J, *et al.* Interobserver, intraobserver and inpatient reliability scores of myocardial strain imaging with 2-d echocardiography in patients treated with anthracyclines. *Ultrasound Med Biol* 2009;35:697-704.
35. Berne RM, Levy MN. *Physiology*. 3rd ed. St. Louis: Mosby; 1993.
36. Voigt JU, Pedrizzetti G, Lysyansky P, *et al.* Definitions for a common standard for 2D speckle tracking echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. *J Am Soc Echocardiogr* 2015;28:183-93.
37. Kuznetsova T, Herbots L, Richart T, *et al.* Left ventricular strain and strain rate in a general population. *Eur Heart J* 2008;29:2014-23.
38. Mele D, Rizzo P, Pollina AV, Fiorencis A, Ferrari R. Cancer therapy-induced cardiotoxicity: role of ultrasound deformation imaging as an aid to early diagnosis. *Ultrasound Med Biol* 2015;41:627-43.
39. Koopman LP, Slorach C, Hui W, *et al.* Comparison between different speckle tracking and color tissue Doppler techniques to measure global and regional myocardial deformation in children. *J Am Soc Echocardiogr* 2010;23:919-28.
40. Nakai H, Takeuchi M, Nishikage T, Lang RM, Otsuji Y. Subclinical left ventricular dysfunction in asymptomatic diabetic patients assessed by two-dimensional speckle tracking echocardiography: correlation with diabetic duration. *Eur J Echocardiogr* 2009;10:926-32.
41. Bates MG, Newman JH, Jakovljevic DG, *et al.* Defining cardiac adaptations and safety of endurance training in patients with m.3243A>G-related mitochondrial disease. *Int J Cardiol* 2013;168:3599-608.
42. Weidemann F, Breunig F, Beer M, *et al.* Improvement of cardiac function during enzyme replacement therapy in patients with Fabry disease: a prospective strain rate imaging study. *Circulation* 2003;108:1299-301.
43. Faber L, Prinz C, Welge D, *et al.* Peak systolic longitudinal strain of the lateral left ventricular wall improves after septal ablation for symptomatic hypertrophic obstructive cardiomyopathy: a follow-up study using speckle tracking echocardiography. *Int J Cardiovasc Imaging* 2011;27:325-33.
44. St John Sutton M, Ky B, Regner SR, *et al.* Longitudinal strain in Friedreich Ataxia: a potential marker for early left ventricular dysfunction. *Echocardiography (Mount Kisco, NY)* 2014;31:50-7.
45. Cheng S, Xanthakis V, Sullivan LM, *et al.* Correlates of echocardiographic indices of cardiac remodeling over the adult life course: longitudinal observations from the Framingham Heart Study. *Circulation* 2010;122:570-8.
46. Ernande L, Rietzschel ER, Bergerot C, *et al.* Impaired myocardial radial function in asymptomatic patients with type 2 diabetes mellitus: a speckle-tracking imaging study. *J Am Soc Echocardiogr* 2010;23:1266-72.

47. Tee M, Noble JA, Bluemke DA. Imaging techniques for cardiac strain and deformation: comparison of echocardiography, cardiac magnetic resonance and cardiac computed tomography. *Expert review of cardiovascular therapy* 2013;11:221-31.
48. Stoodley PW, Richards DA, Hui R, *et al.* Two-dimensional myocardial strain imaging detects changes in left ventricular systolic function immediately after anthracycline chemotherapy. *European journal of echocardiography : the journal of the Working Group on Echocardiography of the European Society of Cardiology* 2011;12:945-952.
49. Pfeffer G, Horvath R, Klopstock T, *et al.* New treatments for mitochondrial disease-no time to drop our standards. *Nature reviews Neurology* 2013;9:474-81.
50. Choi EY, Rosen BD, Fernandes VR, *et al.* Prognostic value of myocardial circumferential strain for incident heart failure and cardiovascular events in asymptomatic individuals: the Multi-Ethnic Study of Atherosclerosis. *Eur Heart J* 2013;34:2354-61.
51. Kalam K, Otahal P, Marwick TH. Prognostic implications of global LV dysfunction: a systematic review and meta-analysis of global longitudinal strain and ejection fraction. *Heart* 2014;23:2014-305538.





# Chapter 10

Serum Fibroblast Growth  
Factor 21 (FGF21) levels in  
adult m.3243A>G carriers:  
clinical implications



## Serum Fibroblast Growth Factor 21 (FGF21) levels in adult m.3243A>G carriers: clinical implications

**Saskia Koene**<sup>1</sup>, Paul de Laat<sup>1</sup>, Doorlène H. van Tienoven<sup>2</sup>, Dennis Vriens<sup>3</sup>, André M. Brandt<sup>2</sup>, Fred C.G.J. Sweep<sup>2</sup>, Richard J.T. Rodenburg<sup>1</sup>, A. Rogier T. Donders<sup>4</sup>, Mirian C.H. Janssen<sup>1,5</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc

<sup>2</sup>Department of Laboratory Medicine, Radboudumc

<sup>3</sup>Department of Radiology and Nuclear Medicine, Radboudumc

<sup>4</sup>Department for Health Evidence, Radboudumc

<sup>5</sup>Department of General Internal Medicine, Radboudumc

Neurology 2014 Jul 8;83(2):125-33

The objective of this study was to determine the value of fibroblast growth factor 21 (FGF21), a recently discovered biomarker for mitochondrial disease, in predicting clinical disease severity and disease progression in adult carriers of the m.3243A>G mutation. We compared the prognostic value of FGF21 with mtDNA heteroplasmy percentages. In the context of a national inventory, the heteroplasmy levels of the m.3243A>G mutation were measured in leukocytes and urinary epithelial cells. The Newcastle Mitochondrial Disease Adult Scale was determined and blood was drawn for measuring FGF21 concentration. This prognostic study included 99 adult carriers of the m.3243A>G mutation. Our analysis revealed a moderate, significant correlation between FGF21 concentration and disease severity ( $r = 0.49$ ;  $p = <0.001$ ). No significant correlations were found between disease severity and the heteroplasmy percentage determined in urinary epithelial cells or the heteroplasmy percentage determined in leukocytes. Weak yet significant correlations were also found between FGF21 concentration and the severity of the myopathy ( $r = 0.38$ ;  $p = <0.001$ ) and between the concentration of FGF21 and the severity of the encephalopathy ( $r = 0.30$ ;  $p = <0.001$ ). Repeated measurements following 25 subjects for two years revealed no significant correlation between FGF21 concentration and disease progression. We conclude that measuring FGF21 concentration had little added value in monitoring and predicting the disease course in this specific patient group.

## Introduction

Within an individual, the mutation load (*i.e.* heteroplasmy) of mitochondrial DNA (mtDNA) mutations can vary widely among tissues. Despite some exceptions,<sup>1</sup> high levels of heteroplasmy are often associated with a measurable deficiency in oxidative phosphorylation and with more severe signs and symptoms.<sup>2</sup> With respect to the m.3243A>G mutation, the most common mutation in mtDNA, high heteroplasmy levels are often associated with a severely debilitating syndrome called mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) syndrome, whereas moderate heteroplasmy levels are associated with milder phenotypes including *e.g.* maternally inherited diabetes and deafness (MIDD) or myopathic symptoms.<sup>3,4</sup> However, because the percentages of heteroplasmy measured in urine epithelial cells and leukocytes are only weakly to moderately correlate with disease severity in m.3243A>G patients,<sup>4,5</sup> using the percentage of heteroplasmy measured in accessible tissues such as blood and urine has limited prognostic value in individual patients. Therefore, a more reliable method is needed in order to predict disease severity of these patient's families.

A recent study reported that serum FGF21 is a novel diagnostic biomarker for differentiating between patients with muscle-manifesting mitochondrial disease and other neuromuscular disorders.<sup>6</sup> In a large cohort of patients with confirmed mitochondrial disease, the sensitivity and specificity of FGF21 was 92.3% and 91.7%, respectively. More recent studies have confirmed that patients with mitochondrial disease have increased FGF21 concentrations,<sup>7,8</sup> although the results obtained from patients and healthy volunteers overlapped.<sup>6,8,9,10</sup> In the present study, we examined the prognostic value of FGF21 in indicating disease severity and monitoring disease progression in adult carriers of the m.3243A>G mutation.

## Methods

### Hypothesis

Our pre-specified hypothesis was that the concentration of fibroblast growth factor 21 (FGF21) in serum correlated with the clinical disease severity, as measured by the Newcastle Mitochondrial Disease Adult Score in adults with carrying the m.3243A>G mutation. We assumed that this correlation would be higher compared to the correlation with heteroplasmy percentages in urinary epithelial cells (UEC) and leukocytes. Furthermore, we hypothesised that an increase in FGF21 would correlate positively with disease progression in 25 randomly selected carriers during a two year follow-up.

## FGF21

We collected blood from patients with a confirmed m.3243A>G mutation and from their maternal relatives; these patients and relatives provided blood samples in the context of the 'National Inventory of Patients with the m.3243A>G mutation in the Netherlands', which was initiated in May 2010.<sup>5</sup> This study investigates the clinical and genetic features of patients with the m.3243A>G mutation and their maternal relatives. To date, 117 adult subjects from 48 families have been enrolled.

The concentration of FGF21 was measured using a commercially available Human FGF21 enzyme-linked immuno sorbent assay (ELISA) kit for detecting human FGF21 (Biovendor, Brno, Czech Republic), following the instruction of the manufacturer. Each sample was measured in duplicate. The inter-assay and intra-assay variability for high and low values was determined based on high and low controls samples, respectively. The functional sensitivity determined at 20% covariance (CV) was 7.35 pg/ml. At a level of 106 pg/ml the within-assay CV was 9.1% and the between assay CV was 9.2%. At a level of 618 pg/ml the within-assay CV was 2.5% and the between-assay CV was 6.4%. Samples with an initial intra-assay CV >15% (in samples with an FGF21 concentration > 100 pg/ml) or >25% (in samples with a first FGF21 concentration <100 pg/ml,) were repeated; if a high intra-assay CV was detected a second time, these samples were excluded from further analysis. According to the kit's manufacturer, the assay has no cross reactivity with human FGF19 and FGF23. The two values obtained from the duplicated measurements were averaged, and the mean of these two measurements was used for further analysis. Samples with a concentration higher than the highest standard value were diluted, and the ELISA was repeated. Reference values (< 200 pg/ml) were obtained from the first study of FGF21 in mitochondrial disease.<sup>6</sup>

Phlebotomy was performed at the outpatient clinic during appointments that were made by the subjects (*i.e.* the appointments occurred at random times of the day during regular office hours). The subjects were not instructed to fast prior to the procedure. Because the phlebotomy was performed immediately following a standardised clinical assessment, the timing of the initial and follow-up appointments was used to calculate the difference in time of the day the samples were taken for the patients included in this follow-up study. No other biomarkers, such as lactate, pyruvate or creatine kinase were evaluated.

Within 24 hours of obtaining the blood samples, the serum was stored in cryotubes at -20°C. In accordance with the kit's manufacturer's instructions, the samples were stored -20°C for a maximum of 10 months and multiple freeze-thaw cycles were avoided. After thawing at 4°C, the samples were analysed within 16 hours. According to the kit's manufacturer, handling the samples in this manner does not result in a decline in the concentration of FGF21 in serum.

## Clinical and genetic assessment

In the 'National Inventory of Patients with the m.3243A>G mutation', all patients and ~50-60% of all maternal relatives (including both symptomatic or asymptomatic relatives) with the m.3243A>G mutation are seen approximately once every two years at our outpatient clinic. All carriers were assessed by the same clinical investigator. Patients were searched for in our genetic database; maternal relatives were recruited for participation by a (maternally related) patient with a proven m.3243A>G mutation. In this study, a total of 135 carriers - 118 of whom were adults - from 48 families have been seen to date. In 78 percent of the carriers included in 2010-2011 (82 adults and 12 children), the study was repeated after two years (76% of the adults; 83% of the children), 3 of the participants (3%; one child) died and 18 subjects (19%; one child) were (temporarily) lost to follow-up (mostly because of insurance or coping issues). Forty-one carriers (5 children) were seen for the first time at the time of the follow-up visit of the first subjects (2012). Twenty-five patients were selected at random from the follow-up group by colleagues who were not involved in the analyses of these results.

The Newcastle Mitochondrial Disease Adult Scale (NMDAS)<sup>11</sup> was used to assess the presence and severity of mitochondrial disease-related clinical symptoms. The NMDAS contains the following four sections: i) current function; ii) system specific involvement; iii) current clinical assessment; and iv) quality of life (QoL). In our analysis, we used sections i-iii to calculate disease severity. Severe mitochondrial disease was defined previously as an NMDAS score above 20.<sup>5</sup> For this study, we defined asymptomatic disease as an NMDAS = 0; mild mitochondrial disease as an NMDAS score of 1 through 5 and moderate mitochondrial disease as an NMDAS score of 6 through 20. We quantified 'myopathy' using the total score for the following items: exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy. 'Encephalopathy' was quantified using the total score obtained from the following items: psychiatric symptoms, migraine, seizure, stroke-like episodes, encephalopathy and cognition. The presence and severity of diabetes mellitus (DM) was obtained from the NMDAS. Renal function was assessed by measuring glomerular filtration rate (using the MDRD equation) and using albumin levels. Microalbuminuria was defined as an albumin-to-creatinine-ratio of >2.0 g/mol for men and >2.5 g/mol for women. Renal failure was defined as a glomerular filtration rate <60 ml/min/1.73m<sup>2</sup>. Patients were classified as having renal failure only, microalbuminuria only, both or neither.

In each subject, the heteroplasmy percentage in the UEC and leukocytes was determined using pyrosequencing,<sup>12</sup> a heteroplasmy percentage ≥5% can be detected using this technique. Subjects with a detectable heteroplasmy percentage in either leukocytes, UEC and/or buccal mucosa cells (the results of buccal mucosa cells are not shown) were considered to be carriers of the mutation. Subjects with heteroplasmy percentages ≤4% (the assay's detection limit) in all tissues were included in the family-matched control group;

note that although the mutation was not detected in three different tissues (leukocytes, UEC and buccal mucosa cells) we cannot be 100% certain that these family members do not carry the mutation. In 7 of these family-matched controls, heteroplasmy percentages were also determined using a more sensitive assay (detection limit 1%); in none of these 7 family-matched controls the mutation was found. The FGF21 concentration in randomly selected non-related healthy controls can be found in published literature.<sup>13</sup>

We also analysed nine paediatric carriers; because the Newcastle Mitochondrial Disease Scale is different for children than for adults, we analysed the correlation between serum FGF21 concentration and disease severity score separately, using the Paediatric version of the Newcastle Mitochondrial Disease Scale in carriers under 18 years of age.

### Standard Protocol Approvals, Registrations, and Patient Consents

This study was approved by the regional Medical Research Ethics Committee. In accordance with the Helsinki agreement, written informed consent was obtained from each participant (or the participant's legal guardian in the case of paediatric subjects) in the context of the 'National Inventory of Patients with the m.3243A>G mutation'.

### Statistics

All parameters were assessed for (log)normality. To prevent non-real values for zero values of the NMDAS and its symptom specific items and heteroplasmy, all values were increased by 1 prior to log-transformation. Variables with a (log)normal distribution, were compared using parametric tests, and the mean and 95%-confidence intervals are reported. Variables that deviated strongly from a (log)normal distribution were analysed by performing a non-parametric test and the median and interquartile ranges are reported. Outliers were not excluded from any analyses. Missing data were not replaced. In case a high number tests were performed (5 or more), critical p-values adjusted using the Bonferroni method (*i.e.*  $0.05/n$ ,  $n$  = number of tests) leading to a critical  $p = 0.0019$  for 26 comparisons and a critical  $p = 0.0013$  for 38 correlations in our manuscript. Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website (<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>; consulted 09-Dec-2013). Thus, a correlation ( $r$ ) of 0.8 to 1.0 is considered a very strong relationship;  $r = 0.6$  to 0.79 is considered a strong relationship;  $r = 0.4$  to 0.59 is considered a moderate relationship;  $r = 0.2$  to 0.39 is considered a weak relationship; and  $r = 0.0$  to 0.19 is considered a weak or no relationship.

The relative difference between the first and the second measurements was calculated by dividing the second measurement by the first measurement, and expressing as the difference as a percentage change. The absolute difference was also calculated by subtracting the first measurement from the second measurement.



Because 40% percent of the variance in concentration of FGF21 is determined genetically,<sup>14</sup> we performed two additional analyses to correct for the effect of kinship in our analyses. First, we performed a separate analysis that included only the most severely affected patient in each family. Second, we used generalised estimating equation models (working correlation structure: independence, with robust standard error for correction) to predict the contribution of covariates and candidate predictors with correction for kinship. Gender, age, body mass index (BMI), heteroplasmy percentage in UEC and leukocytes, myopathy, encephalopathy, DM, renal function classification and the disease severity (NMDAS) were included as possible covariates and candidate predictors for FGF21. Gender, age, BMI, heteroplasmy percentage in UEC and leukocytes and the concentration of FGF21 were included as possible covariates and candidates for disease severity (NMDAS score).

Because genetic factors between family members are not likely to influence intra-individual changes in FGF21 levels over time, the influence of covariates in the changes during follow-up was calculated using linear regression models. The changes in BMI, the change in sampling time and the changes in myopathy, encephalopathy, diabetes mellitus and disease severity were included as covariates and candidate predictors for the change in FGF21 concentration. The changes in BMI, heteroplasmy percentage in UEC and leukocytes, and FGF21 concentration were included as covariates and candidate predictors for the change in disease severity (NMDAS severity).

All analyses were performed using IBM's SPSS statistics software packages, version 20.0.0.

**Table 1.**  
**Patient characteristics of the cohort of all adult m.3243A>G carriers and the follow-up cohort.** Patient characteristics for all adult carriers (*n* = 99) and for the follow-up patients (*n* = 25). The change in the 25 patients in the follow-up group was calculated as the ratio of the second measurement and the first measurement and is expressed as the percentage that the variable changed relative to the first measurement.

All carriers (n = 99)			Follow-up patients (n = 25)			All carriers compared to follow-up patients		Change (% higher)	
Parameter	Median (IQR)	Range	n =	Mean (95%CI)	Range	n =	p =	Median change (IQR)	Range
Gender (% female)	72		99	64		25	0.56		
Age (years)	46 (34 - 56)	18 - 81	99	44 (20 - 68)	19 - 64	25	0.31		
BMI (kg/m <sup>2</sup> )	23.8 (20.8 - 26.6)	16.8 - 40.9	64	22.4 (16 - 31.4)	17.6 - 35.8	12	0.64	1b (-5 - +7d)	+0 - + 10
Heteroplasmy percent- age leucocytes (%)	17 (8 - 26)	2 - 56	98	22 (0 - 48)	2 - 51	25	0.02	0 (-15 - + 8)	-50 - + 133
Heteroplasmy percent- age UFC (%)	46 (28 - 70)	4 - 96	96	55 (4.8 - 105)	8 - 98	23	0.007	0 (-8 - + 10)	-37 - + 350
NMDAS score	14 (6 - 22)	0 - 82	99	16a (7 - 26c)	0 - 74	25	0.07	9 (-13 - + 30)	-50 - + 400
Domain 1	5 (1 - 9)	0 - 38	99	8a (3.5 - 12)	0 - 35	25	0.05	0 (-40 - + 33)	-99 - + 200
Domain 2	5 (2 - 9)	0 - 24	99	6a (2 - 9.5)	0 - 19	25	0.03	0 (-24 - + 25)	-100 - + 100
Domain 3	5 (2 - 9)	0 - 20	99	3 (0 - 7)	0 - 20	25	0.39	0 (-75 - + 20)	-100 - + 100
Myopathy score	2 (1 - 4)	0 - 14	99	2 (1 - 4)	0 - 11	25	0.19	0 (-50 - + 38)	-100 - + 100
Encephalopathy score	2 (0 - 3)	0 - 20	99	1 (1 - 3)	0 - 22	25	0.73	-26 (-100 - + 100)	-100 - + 100
DM severity	0 (0 - 5)	0 - 5	99	1.8 (0 - 6.4)	0 - 5	25	0.60	-34 (-99 - + 0)	-100 - + 125
FGF21 (pg/ml)	263 (140 - 523)	3 - 3,742	99	434 (41 - 1,870)	56 - 1,776	25	0.28	+36 (-26 - + 63)	-73 - + 7,025

BMI = body mass index; CI = 95% confidence interval; DM = diabetes mellitus; Encephalopathy score = sum of the encephalopathic symptoms of the NMDAS (psychiatric symptoms, migraine, seizure, stroke-like episodes, encephalopathy and cognition); FGF21 = Fibroblast Growth Factor 21; IQR = interquartile range; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); *n* = number of carriers of which data were available at that specific time point; NMDAS = Newcastle Mitochondrial Disease Adult Scale; Section 1 = current function; Section 2 = system specific involvement; Section 3 = current clinical assessment; UEC = urinary epithelial cells; \*signifi-  
cantly different (*p* < 0.0019), a = median is given instead of mean; b = mean is given instead of median; c = IQR is given instead of CI; d = CI is given instead of IQR

## Results

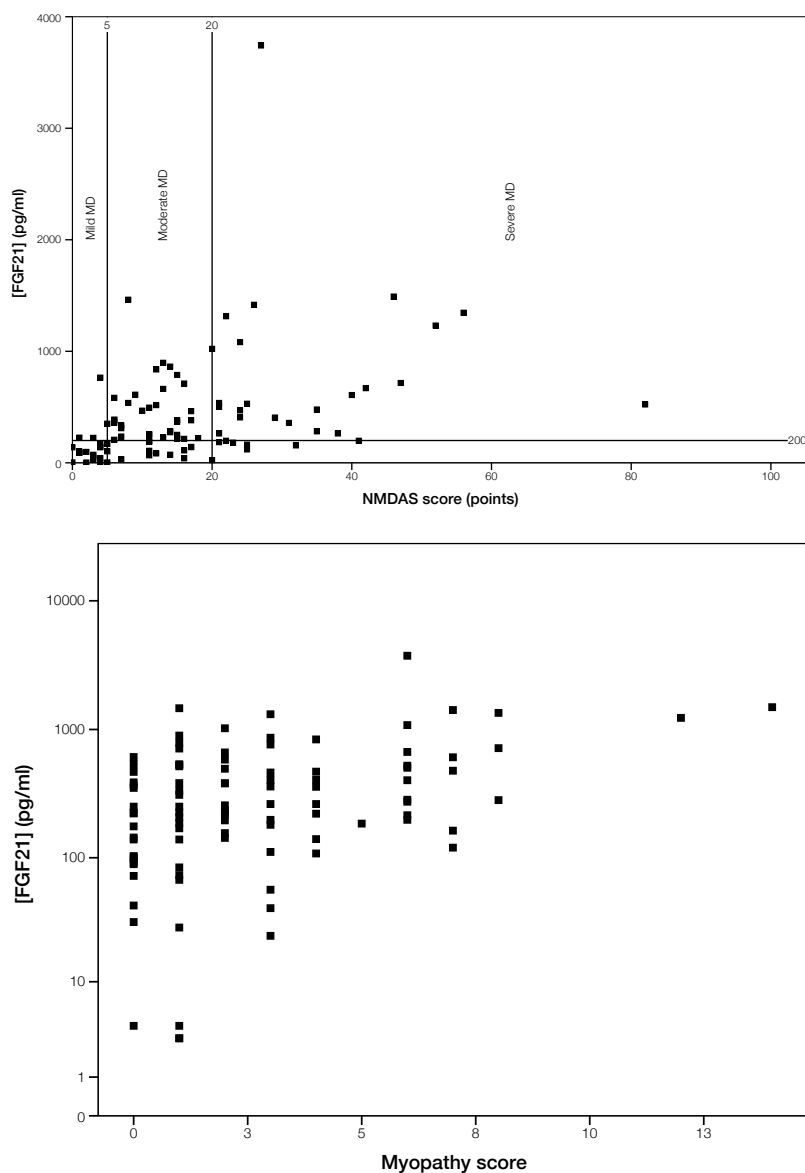
### Patient characteristics

We initially included 118 adult subjects in our study. For a variety of reasons (e.g., the samples were not available or were insufficient for measuring FGF21), we were unable to determine the FGF21 levels in 15 of these subjects. Three additional subjects were excluded because of high intra-assay covariance. Thus, our final cohort included 99 adult carriers of the m.3243A>G mutation. These carriers were from 39 distinct families (median: 3 subjects per family; range: 1-10 subjects per family; Table 1). This cohort contained two asymptomatic patients (2%), 20 patients with mild mitochondrial disease (20%), 47 patients with moderate mitochondrial disease (48%), and 30 patients with severe mitochondrial disease (30%). Since not always sufficient material was available, heteroplasmy percentage in UEC was available for 96 subjects; heteroplasmy percentage in leukocytes was available for 98 patients.

At the time of our analysis, renal function had been studied in 88% of the subjects. Four patients had renal problems (with FGF21 levels of 403, 236, 21 and 16 pg/ml). Three other patients had received a kidney transplant; two of these patients had a normal transplant function (with FGF21 levels of 407 and 1,082 pg/ml) and the third patients had a moderate transplant function (with an FGF21 level of 663 pg/ml). Finally, one severely affected patient had focal segmental glomerulosclerosis, but normal renal function (with an FGF21 level of 3,742 pg/ml).

### The value of FGF21 as an indicator of clinical disease severity

In the total cohort containing 99 subjects, the correlation between the concentration of FGF21 and total NMDAS score was  $r = 0.49$  ( $p = <0.001$ ; Figure 1A and Table 2). The three outliers in Figure 1A represent a 47-year-old female with diabetes mellitus, hearing loss, focal segmental glomerulosclerosis, cardiomyopathy, exercise intolerance and mild cognitive disturbances (FGF21 = 3,742 pg/ml), an 18-year-old asymptomatic male with a heteroplasmy percentage of 93% in the urine (FGF21 = 523 pg/ml) and a 31-year-old female with exercise intolerance and clinical obesity (FGF21 = 1,463 pg/ml). No significant correlation was found between the heteroplasmy percentage in UEC and the NMDAS score, nor between the heteroplasmy percentage in leukocytes and the NMDAS score (Table 2). Among the 47 patients with mild or moderate disease, no significant correlation was found between the total NMDAS score and FGF21 concentration or heteroplasmy percentages (either measured in UEC or leukocytes). Among the 30 patients with severe disease, no significant correlation was found between the total NMDAS score and the heteroplasmy percentage in either UEC or leukocytes.



**Figure 1.**

**Correlation between the concentration of FGF21 and disease severity/ and myopathy score in adult carriers of the m.3243A>G mutation.** A. The correlation between the concentration of FGF21 and general disease severity (total NMDAS) is  $r = 0.49$  ( $p < 0.001$ ). The interquartile range is 140 - 523 pg/ml. The two vertical lines represent the NMDAS score reference values for mild (<5) , moderate (5-19) and severe (≥20) mitochondrial disease. The horizontal line represents the FGF21 reference value (≤200 pg/ml). B. The correlation between the concentration of FGF21 and the myopathy score is  $r = 0.38$  ( $p < 0.001$ ). The y-axis is loglinear. FGF21 = Fibroblast Growth Factor 21; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); MD = Mitochondrial Disease; NMDAS = Newcastle Mitochondrial Disease Adult Scale

Because FGF21 was described previously as a marker for muscle-manifesting mitochondrial disease,<sup>6</sup> we also examined the correlation between FGF21 concentration and the muscle symptoms rated by the NMDAS (*i.e.* the myopathy score). We found a weak but significant correlation between the myopathy score and FGF21 concentration ( $r = 0.38$ ;  $p = <0.001$ ). No significant correlation was found between myopathy score and heteroplasmy percentage in leukocytes or UEC (Table 2).

**Table 2.**

**Correlations between NMDAS score and myopathy score and FGF21 and the heteroplasmy percentages in leukocytes and UEC at baseline and at follow-up.** Correlations between the total NMDAS score and the myopathy score and the concentration of FGF21 and the heteroplasmy percentages for all adult carriers ( $n = 99$ ) and between the ratios (second measurement/first measurement) of these parameters for the follow-up patients ( $n = 25$ ). \*Significant correlations ( $p < 0.0013$ ).

	Baseline ( $n = 99$ )		Follow-up ( $n = 25$ )	
	Correlation with NMDAS ( $r$ )	Correlation with myopathy score ( $r$ )	Correlation between ratio's and the ratio of the NMDAS ( $r$ )	Correlation between ratio's and the ratio of the myopathy score ( $r$ )
FGF21 in serum (pg/ml)	<b>0.49 (<math>p = &lt;0.001</math>)</b>	<b>0.38 (<math>p = &lt;0.001</math>)</b>	0.35 ( $p = 0.09$ )	0.47 ( $p = 0.02$ )
Heteroplasmy in leukocytes (%)	0.15 ( $p = 0.13$ )	0.21 ( $p = 0.43$ )	-0.15 ( $p = 0.48$ )	0.39 ( $p = 0.06$ )
Heteroplasmy in UEC (%)	0.29 ( $p = 0.03$ )	0.19 ( $p = 0.68$ )	-0.32 ( $p = 0.14$ )	0.31 ( $p = 0.14$ )

\*Significant correlations ( $p < 0.0013$ ); FGF21 = Fibroblast Growth Factor 21; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); NMDAS = Newcastle Mitochondrial Disease Adult Scale; UEC = urinary epithelial cells

Consistent with previous reports,<sup>15</sup> we also found a negative correlation between age and the heteroplasmy percentage in leukocytes ( $r = -0.56$ ;  $p < 0.001$ ;  $n = 98$ ). A weak negative correlation was also found between age and the heteroplasmy percentage in UEC ( $r = -0.23$ ;  $p < 0.001$ ;  $n = 96$ ). Finally, a moderate positive correlation was found between age and FGF21 concentration ( $r = 0.49$ ;  $p < 0.001$ ).

Correcting for kinship, generalised estimating equations revealed a significant contribution of the age at assessment to the disease severity (NMDAS) ( $r = 0.05$ ;  $n = 62$  (63%)), but gender, BMI, FGF21 concentration and heteroplasmy percentage in leukocytes or UEC did not contribute significantly to disease severity. We found a moderate positive correlation between age and the NMDAS score ( $r = 0.40$ ;  $p = <0.001$ ). Correcting for kinship, generalised estimating equations, revealed no significant contribution of age at assessment, gender, BMI, FGF21 concentration, to heteroplasmy percentage in leukocytes or UEC.

Among the 37 mutation carriers with normal levels of FGF21 ( $\leq 200$  pg/ml), two subjects were asymptomatic (5.4%) and 16 subjects had mild mitochondrial disease (44%). Carriers with elevated FGF21 levels ( $>200$  pg/ml) had higher scores on the first section of the NMDAS, however, their heteroplasmy percentages in leukocytes and UEC, total NMDAS scores, myopathy and encephalopathy scores, DM score and BMI values were similar to the carriers with normal FGF21 levels. The FGF21 levels these two cohorts are summarised in Table 3.

**Table 3.**

**Patient characteristics of m.3243A>G carriers with normal concentrations of FGF21 versus carriers with increased levels of FGF21.** The characteristics of carriers with a normal concentration of FGF21 ( $\leq 200$  pg/ml;  $n = 37$ ) and those with increased levels of FGF21 ( $>200$  pg/ml;  $n = 62$ ) are shown. Medians are given. For the categorical disease severity, the percentages of the total group are given as well.

	[FGF21] normal ( $n = 37$ )	[FGF21] increased ( $n = 62$ )
BMI (kg/m <sup>2</sup> )	23.4	24.4
Heteroplasmy leukocytes (%)	12	19*
Heteroplasmy UEC (%)	36	52*
Asymptomatic	2 (5%)	0
Mild disease	16 (43%)	4 (6%)
Moderate disease	11 (30%)	34 (55%)
Severe disease	3 (22%)	24 (39%)
Total NMDAS score	7	15*
Domain 1	1	6*
Domain 2	4	6*
Domain 3	2	4*
DM score	0	0
Encephalopathy score	1	2*
Myopathy score	1	2*

\*significantly higher ( $p < 0.0019$ ); BMI = body mass index; DM = diabetes mellitus; Section 1 = current function; Section 2 = system specific involvement; Section 3 = current clinical assessment; Encephalopathy score = sum of the encephalopathic symptoms of the NMDAS (psychiatric symptoms, migraine, seizure, stroke-like episodes, encephalopathy and cognition); FGF21 = Fibroblast Growth Factor 21; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); NMDAS = Newcastle Mitochondrial Disease Adult Scale; UEC = urinary epithelial cells

Correcting for kinship, generalised estimating equations revealed no significant contribution of gender, age, BMI, heteroplasmy percentage in UEC or leukocytes, myopathy, encephalopathy, DM, renal function and the disease severity (NMDAS score) to FGF21 concentration.

One patient died at the age 58 years during the year of follow-up period due to heart failure secondary to m.3243A>G-related hypertrophic cardiomyopathy. The concentration of FGF21 in this patient at the time of assessment was 1,231 pg/ml. The 3-year old child who died during follow-up due to intractable epilepsy, had an FGF21 concentration of 21 pg/ml one year prior to his death.

We next analysed only the most severely affected patient in each of the 39 families. The results of this subset analysis were similar to those found in the entire cohort. Specifically, the correlation between the total NMDAS score and FGF21 concentration was 0.56 ( $p = <0.001$ ). No significant correlation was found between myopathy score and FGF21 concentration ( $r = 0.48$ ;  $p = 0.002$ ). Moreover, no significant correlation was found between the heteroplasmy percentage in UEC and the NMDAS score and/or the myopathy score.

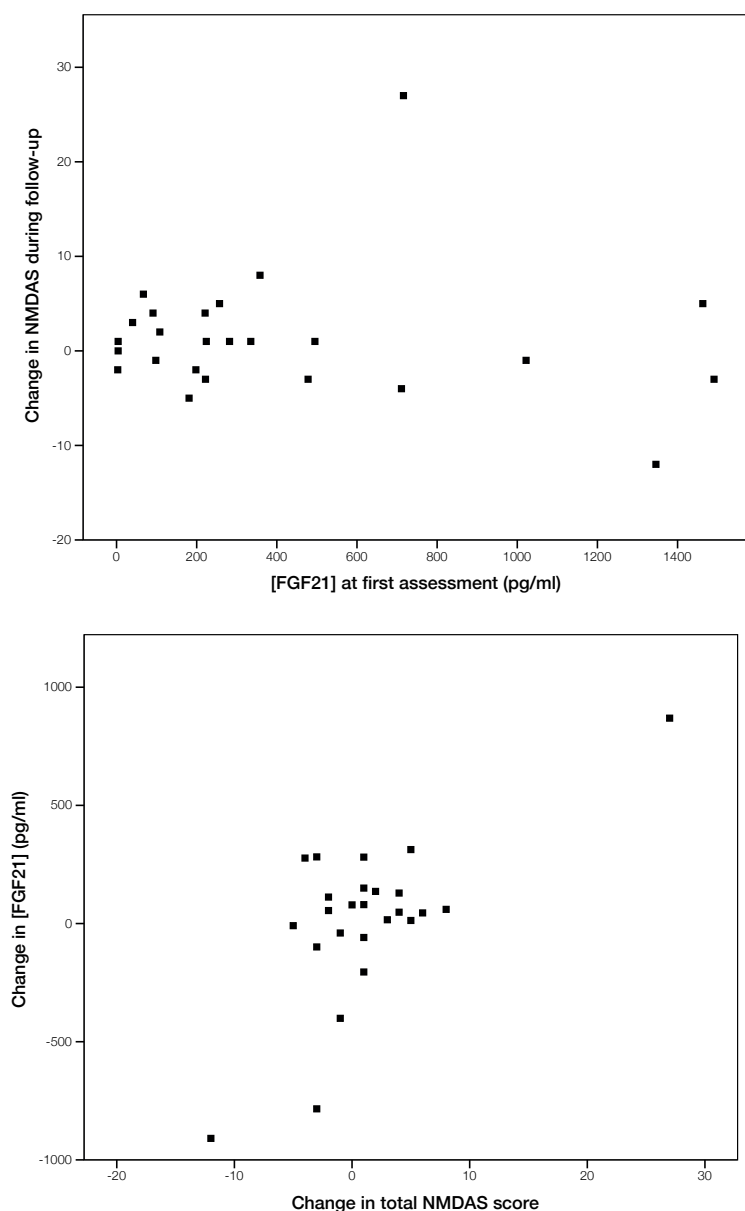
### Follow-up data

In a subset of 25 carriers who were selected at random from the initial cohort of 99 subjects, FGF21 concentration and NMDAS scores were measured approximately two years after the initial assessment. The median interval between the two assessments was 20 months (range 15 - 24 months). These 25 carriers in the follow-up study were similar to the full cohort with respect to gender, age, BMI, FGF21 concentration, heteroplasmy percentages and NMDAS total and subsections scores. With respect to time of the day, the follow-up sample was obtained 1.5 hours later in the day (median value; range 0.5 hours earlier up to 3 hours later) than the first sample. The characteristics of the entire cohort and the follow-up subgroup of subjects are summarised in Table 1.

### The value of FGF21 in predicting clinical disease progression

We found no significant correlation between the difference in NMDAS score (*i.e.* the change in NMDAS score from the follow-up visit to the second follow-up visit) and the difference in FGF21 concentration (*i.e.* the change in FGF21 concentration from the first follow-up to the second follow-up visit; Figure 2B and Table 2). Similar results were obtained with respect to the change in the heteroplasmy percentages in leukocytes and UEC between the first and second follow-up visits.

We found no correlation between the change in the myopathy score and the change in the concentration of FGF21 ( $r = 0.47$ ;  $p = 0.017$ ;  $n = 25$ ). There was no significant correlation between myopathy progression and the change in the heteroplasmy percentage in leukocytes or UEC (Table 2).



**Figure 2.**

**Predictive and monitoring value of FGF21 in estimating disease severity in adult carriers of the m.3243A>G mutation** A) There is no significant correlation ( $r = -0.11$ ;  $p = 0.601$ ) between the concentration of FGF21 at the first assessment and the progression of the disease (here displayed as the difference between the second measurement and the first measurement). B) There is no significant correlation ( $r = 0.35$ ;  $p = 0.086$ ) between the ratio of the first and the second FGF21 concentration measurements and the ratio of the first and the second NMDAS measurements. The outliers are discussed in the text. FGF21 = Fibroblast Growth Factor 21; NMDAS = Newcastle Mitochondrial Disease Adult Scale



No significant correlation was found between FGF21 concentration at the first follow-up visit and disease progression during follow-up ( $r = -0.11$ ;  $p = 0.60$ ). Moreover, no significant correlation was found between the disease progression during follow-up and the initial heteroplasmy percentage measured in leucocytes ( $r = -0.19$ ;  $p = 0.36$ ) or UEC ( $r = 0.035$ ;  $p = 0.87$ ), respectively. The three outliers in Figure 2B are a 32-year-old female with MELAS syndrome and a severely progressive disease with severe depression requiring hospitalisation (the first and second FGF21 levels were 716 and 1,585 pg/ml respectively) a 47-year-old female with MELAS syndrome and a relatively stable disease course (the first and second FGF21 levels were 1,494 and 707 pg/ml respectively) and a 20-year-old man with MELAS syndrome and cardiomyopathy with improving functional abilities in daily life but then developed diabetes mellitus that required insulin treatment (the first and second FGF21 levels were 1,346 and 437 pg/ml respectively).

A linear regression model revealed that neither the change in BMI, the change in heteroplasmy percentages in UEC and leukocytes, nor the change in FGF21 concentration contributed significantly to the change in disease severity (*i.e.* NMDAS score).

A linear regression model also revealed that neither the change in BMI, the change in sampling time nor the change in myopathy, encephalopathy, DM, or disease severity contributed significantly to the change in FGF21 concentration.

### Family matched controls

The family-matched controls were similar to their mutation-carrying maternal relatives with respect to age ( $p = 0.069$ ), gender ( $p = 0.81$ ) and BMI ( $p = 0.14$ ), and encephalopathy scores ( $p = 0.002$ ). As expected, compared to the mutation-carrying relatives, the controls scored significantly lower with respect to their total NMDAS scores ( $p < 0.001$ ) and myopathy scores ( $p < 0.001$ ). The median FGF21 concentration in the family-matched control group was 31 pg/ml (IQR 7 - 173 pg/ml; range 7 - 1,072 pg/ml), which was significantly lower than their mutation-carrying maternal relatives ( $p < 0.001$ ).

### Results in children carrying the m.3243A>G mutation

We also studied nine children, who were carriers of the m.3243A>G mutation, and found no significant correlation between the disease severity and FGF21 concentration.

## Discussion

There is currently an urgent need for biomarkers that can reliably, objectively and quantitatively monitor the effect of new treatments and interventions on the disease course in patients with mitochondrial disease. Measuring the concentration of FGF21, a relatively new diagnostic biomarker, was recently reported to be indicative of clinical severity and disease progression in a heterogeneous cohort of patients with mitochondrial disease.<sup>6</sup> To investigate further the potential prognostic value of this new mitochondrial biomarker with respect to monitoring disease severity and predicting disease progression, we measured serum FGF21 levels in a genetically homogeneous but clinically heterogeneous disease cohort comprised of m.3243A>G carriers, with various levels of heteroplasmy.

In this prognostic study, we first evaluated whether FGF21 concentration can be used to indicate disease severity at a single time point. In clinical practice, heteroplasmy levels are routinely used to cautiously predict the course of the disease, even though these parameters correlate only weakly, or at best moderately, with disease severity in patients carrying the m.3243A>G mutation.<sup>5</sup> In a cohort of 99 carriers of the m.3243A>G mutation, we found that serum FGF21 concentration correlated only moderately with disease severity (measured using the multi-dimensional and detailed NMDAS). Although the correlation between disease severity and FGF21 concentration was stronger compared to the heteroplasmy percentages in leukocytes and urinary epithelial cells, our results suggest that FGF21 level is not suitable as an indicator of disease severity in adult m.3234A>G carriers.

Given the urgent need for a reliable prognostic biomarker for following treatment effects, one might be tempted to take a biomarker that seems suitable for diagnostic purposes and use it for longitudinal follow-up purposes. Based on this reasoning, we systematically studied whether FGF21 could be used to monitor disease progression in 25 of the patients in our study. We found no significant correlation between disease progression and FGF21 concentration (*i.e.* the concentration prior to disease progression and the change in concentration between the two follow-up measurements) during 2-year follow-up. These results suggest that FGF21 may not be a suitable prognostic biomarker for follow-up studies in this particular patient group. Nevertheless, FGF21 could still be studied as a potential biomarker of other mitochondrial diseases.

Because FGF21 has been reported previously as a diagnostic biomarker for muscle-manifesting disease,<sup>6</sup> we also examined the correlation between myopathic clinical features and FGF21 concentration (Table 2). We found a weak correlation between FGF21 concentration and myopathy (*i.e.* myopathy score). In contrast, we found no significant correlation between FGF21 concentration and muscle disease progression nor a

significant correlation between heteroplasmy percentage and myopathy severity. It is important to note that because we did not analyse muscle biopsies (for example, using cytochrome c oxidase staining), and we could only examine the correlation between FGF21 and clinical data.

We also examined the contribution of several previously identified covariates to FGF21 concentration. Because 40% of the variation in FGF21 concentration is genetically determined,<sup>14</sup> we corrected our data for inherited factors other than the m.3243A>G mutation using two methods: first, we analysed only the most severely affected patient in each family, and second, we corrected for kinship when assessing the effect of covariates. The correlation between FGF21 concentration and disease severity strengthened slightly (to  $r = 0.56$  compared to  $r = 0.49$ ) when only one patient per family was included. The following covariate analyses were all corrected for kinship. Firstly, we did not observe the covariance by BMI and by diabetes mellitus, parameters known to correlate with FGF21 levels in large studies.<sup>14-22</sup> However, the pathophysiology of mitochondrial diabetes mellitus seems to involve impaired insulin secretion rather than hyperinsulinism.<sup>23</sup> Secondly, the increased levels of FGF21 that have been described in patients with chronic kidney disease<sup>24</sup> were not observed in seven patients with renal failure and microalbuminuria. Although our study failed to replicate the correlations reported in studies with larger sample sizes, our results suggest that none of these parameters is likely to have caused any major bias with respect to our correlations.

In healthy volunteers, circulating FGF21 levels do not follow a clear diurnal rhythm, and ketosis induced by either a 2-day fast or a ketogenic diet does not significantly increase serum FGF21 levels.<sup>25,26</sup> The median FGF21 level in healthy individuals was 156 pg/ml (90%CI 29 – 1,844 pg/ml) and previous reports showed that this level varies physiologically by 64 - 250 fold between-subjects<sup>16,18</sup> and by approximately 20% within a given subject during the day.<sup>18</sup> Although we observed higher concentrations of FGF21 in the cohort carrying the m.3243A>G mutation compared to family matched controls with undetectable heteroplasmy percentages in three different tissues, the variability in FGF21 levels within the healthy population may complicate the interpretation of FGF21 values in diagnostic and prognostic studies.

A strength of our study is the high number of carriers with different levels of the same mutation in their mtDNA, resulting in a heterogeneous multi-system disease that we quantified systematically. Because we included both patients and dormant (*i.e.* asymptomatic) carriers, we were able to study the role of FGF21 in reflecting the clinical expression of the disease in a representative, heterogeneous population. Studying a genetically homogeneous population reduces potential confounding by individual biochemical abnormalities. Although no clear differences in FGF21 concentration were reported between patients with nuclear mutations and mtDNA mutations (or between

various mtDNA mutations),<sup>6</sup> the ability to generalise these results into other populations with mitochondrial disease warrants further study. Our quantitative follow-up of 25 patients, including both clinical and biochemical parameters enabled us to draw tentative conclusions regarding the value of FGF21 as a prognostic biomarker for monitoring disease progression.

There is currently no gold standard for measuring disease severity in patients with mitochondrial disease. In this study, we used the NMDAS, a multi-dimensional, standardised questionnaire to measure disease severity in patients with mitochondrial disease. It is important to note that the first and the second sections of the NMDAS are obtained from the patient history, thus measuring the most important factor of disease severity (namely the complaints experienced by the patient); however this approach potentially limits the reliability of this parameter. Moreover, the specificity of some of the NMDAS items is relatively low, including frequently commonly reported complaints such as gastrointestinal - and mood problems, as well as complaints common to the ageing process, including hearing loss and exercise intolerance. These aging-related complaints might explain the significant positive contribution of age to the disease severity.

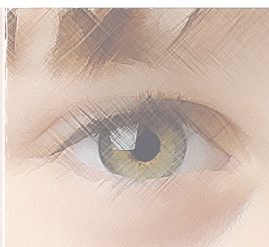
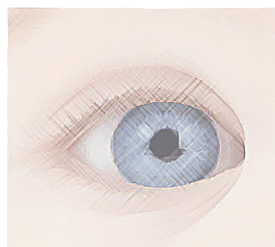
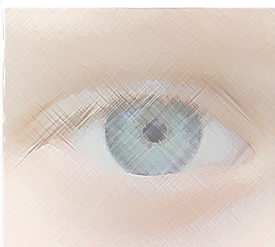
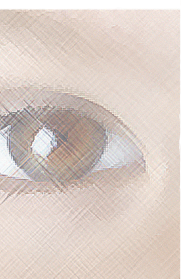
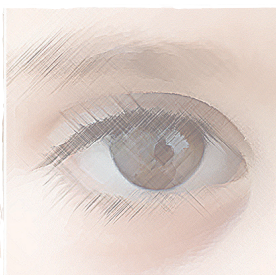
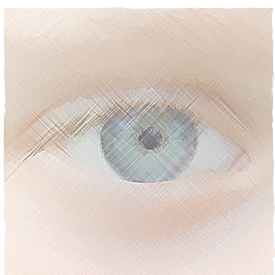
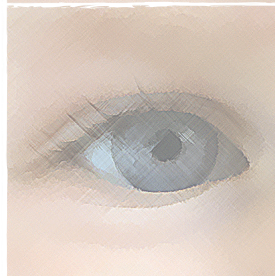
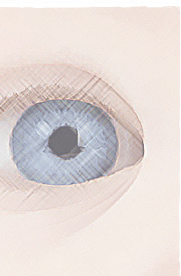
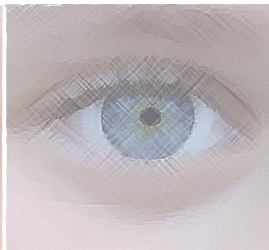
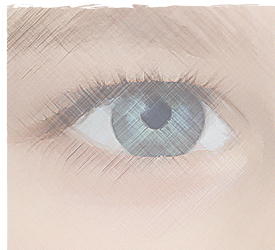
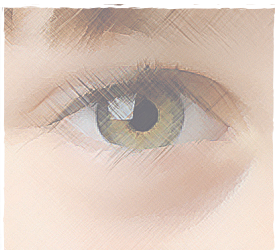
In conclusion, we report that FGF21 concentration correlates moderately with disease severity but not to disease progression in carriers of the m.3243A>G mutation. Therefore, we conclude that FGF21 has little added value in monitoring and predicting the disease course in this specific patient group. In the quest to identify an objective, quantitative, sensitive and responsive prognostic biomarker that correlates with clinically relevant parameters, the validation of functional tests will take the highest priority in the preparation of future clinical trials.

## References

1. Ruiter EM, Siers MH, van den Elzen C, *et al.* The mitochondrial 13513G > A mutation is most frequent in Leigh syndrome combined with reduced complex I activity, optic atrophy and/or Wolff-Parkinson-White. *Eur J Hum Genet* 2007;15:155-161.
2. Janssen AJ, Schuelke M, Smeitink JA, *et al.* Muscle 3243A-->G mutation load and capacity of the mitochondrial energy-generating system. *Annals of neurology* 2008;63:473-481.
3. Sproule DM, Kaufmann P. Mitochondrial encephalopathy, lactic acidosis, and strokelike episodes: basic concepts, clinical phenotype, and therapeutic management of MELAS syndrome. *Annals of the New York Academy of Sciences* 2008;1142:133-158.
4. Laloi-Michelin M, Meas T, Ambonville C, *et al.* The clinical variability of maternally inherited diabetes and deafness is associated with the degree of heteroplasmy in blood leukocytes. *J Clin Endocrinol Metab* 2009;94:3025-3030.
5. de Laat P, Koene S, van den Heuvel LP, Rodenburg RJ, Janssen MC, Smeitink JA. Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A > G mutation. *J Inherit Metab Dis.* 2012 Nov;35(6):1059-1069.
6. Suomalainen A, Elo JM, Pietilainen KH, *et al.* FGF21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. *Lancet Neurol* 2011;10:806-818.
7. Crooks DR, Natarajan TG, Jeong SY, *et al.* Elevated FGF21 secretion, PGC-1alpha and ketogenic enzyme expression are hallmarks of iron-sulfur cluster depletion in human skeletal muscle. *Hum Mol Genet* 2014;1;23;1:24-39.
8. Davis RL, Liang C, Edema-Hildebrand F, *et al.* Fibroblast Growth Factor 21 is a sensitive biomarker of mitochondrial disease. *Neurology* 2013;19;81:1819-1826.
9. Salehi MH, Kamalidehghan B, Houshmand M, Aryani O, Sadeghizadeh M, Mossalaeie MM. Association of fibroblast growth factor (FGF21) as a biomarker with primary mitochondrial disorders, but not with secondary mitochondrial disorders (Friedreich Ataxia). *Mol Biol Rep* 2013;40;11:6495-6499.
10. Su SL, Wang WF, Wu SL, *et al.* FGF21 in ataxia patients with spinocerebellar atrophy and mitochondrial disease. *Clinica chimica acta; international journal of clinical chemistry* 2012;414:225-227.
11. Schaefer AM, Phoenix C, Elson JL, McFarland R, Chinnery PF, Turnbull DM. Mitochondrial disease in adults: a scale to monitor progression and treatment. *Neurology* 2006;66:1932-1934.
12. Lowik MM, Hol FA, Steenbergen EJ, Wetzels JF, van den Heuvel LP. Mitochondrial tRNA<sup>Leu</sup>(UUR) mutation in a patient with steroid-resistant nephrotic syndrome and focal segmental glomerulosclerosis. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* 2005;20:336-341.
13. Angelin B, Larsson TE, Rudling M. Circulating fibroblast growth factors as metabolic regulators--a critical appraisal. *Cell Metab* 2012;16:693-705.
14. Tyynismaa H, Raivio T, Hakkarainen A, *et al.* Liver fat but not other adiposity measures influence circulating FGF21 levels in healthy young adult twins. *J Clin Endocrinol Metab* 2011;96:E351-355.
15. Rahman S, Poulton J, Marchington D, Suomalainen A. Decrease of 3243 A-->G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *J Hum Genet* 2001;68:238-240.
16. Fazeli PK, Misra M, Goldstein M, Miller KK, Klibanski A. Fibroblast growth factor-21 may mediate growth hormone resistance in anorexia nervosa. *J Clin Endocrinol Metab* 2010;95:369-374.

17. Dostalova I, Kavalkova P, Haluzikova D, *et al.* Plasma concentrations of fibroblast growth factors 19 and 21 in patients with anorexia nervosa. *J Clin Endocrinol Metab* 2008;93:3627-3632.
18. Galman C, Lundasen T, Kharitononkov A, *et al.* The circulating metabolic regulator FGF21 is induced by prolonged fasting and PPARalpha activation in man. *Cell metabolism* 2008;8:169-174.
19. Hojman P, Pedersen M, Nielsen AR, *et al.* Fibroblast growth factor-21 is induced in human skeletal muscles by hyperinsulinemia. *Diabetes* 2009;58:2797-2801.
20. Mashili FL, Austin RL, Deshmukh AS, *et al.* Direct effects of FGF21 on glucose uptake in human skeletal muscle: implications for type 2 diabetes and obesity. *Diabetes Metab Res Rev* 2011;27:286-297.
21. Domingo P, Gallego-Escuredo JM, Domingo JC, *et al.* Serum FGF21 levels are elevated in association with lipodystrophy, insulin resistance and biomarkers of liver injury in HIV-1-infected patients. *AIDS*. 2010 Nov 13;24(17):2629-2637.
22. Semba RD, Sun K, Egan JM, Crasto C, Carlson OD, Ferrucci L. Relationship of Serum Fibroblast Growth Factor 21 with Abnormal Glucose Metabolism and Insulin Resistance: The Baltimore Longitudinal Study of Aging. *J Clin Endocrinol Metab*. 2012 Apr;97(4):1375-1382.
23. Hosszufalusi N, Karcagi V, Horvath R, *et al.* A detailed investigation of maternally inherited diabetes and deafness (MIDD) including clinical characteristics, C-peptide secretion, HLA-DR and -DQ status and autoantibody pattern. *Diabetes Metab Res Rev* 2009;25:127-135.
24. Lin Z, Zhou Z, Liu Y, *et al.* Circulating FGF21 levels are progressively increased from the early to end stages of chronic kidney diseases and are associated with renal function in Chinese. *PLoS one* 2011;6:e18398.
25. Lee P, Brychta RJ, Linderman J, Smith S, Chen KY, Celi FS. Mild cold exposure modulates Fibroblast Growth Factor 21 (FGF21) diurnal rhythm in humans: relationship between FGF21 levels, lipolysis, and cold-induced thermogenesis. *J Clin Endocrinol Metab* 2013;98:E98-102.
26. Foo JP, Aronis KN, Chamberland JP, Paruthi J, Moon HS, Mantzoros CS. Fibroblast Growth Factor 21 levels in young healthy females display day and night variations and are increased in response to short-term energy deprivation through a leptin-independent pathway. *Diabetes care* 2013;36:935-942.







# Chapter 11

Serum GDF15 levels  
correlate to mitochon-  
drial disease severity and  
myocardial strain, but not to  
disease progression in adult  
m.3243A>G carriers



## Serum GDF15 levels correlate to mitochondrial disease severity and myocardial strain, but not to disease progression in adult m.3243A>G carriers

**Saskia Koene**<sup>1</sup>, Paul de Laat<sup>1</sup>, Doorlène H. van Tienoven<sup>2</sup>, Gert Weijers<sup>3</sup>, Dennis Vriens<sup>3</sup>, Fred C.G.J. Sweep<sup>2</sup>, Janneke Timmermans<sup>4</sup>, Livia Kapusta<sup>5,6</sup>, Mirian C.H. Janssen<sup>1,7</sup>, Jan A.M. Smeitink<sup>1</sup>

<sup>1</sup>Nijmegen Centre for Mitochondrial Disorders, Radboudumc

<sup>2</sup>Department of Laboratory Medicine, Radboudumc

<sup>3</sup>Department of Radiology and Nuclear Medicine, Radboudumc

<sup>4</sup>Department of Cardiology, Radboudumc

<sup>5</sup>Department of Paediatric Cardiology, Amalia Children's Hospital, Radboudumc

<sup>6</sup>Department of Paediatrics, Paediatric Cardiology Unit, Tel-Aviv Sourasky Medical Centre, Tel Aviv, Israel

<sup>7</sup>Department of General Internal Medicine, Radboudumc

Journal of Inherited Metabolic Diseases – Reports 2015;24:69-81

In this observational cohort study, we examined the prognostic value of growth and differentiation factor 15 (GDF15) in indicating and monitoring general mitochondrial disease severity and progression in adult carriers of the m.3243A>G mutation.

Ninety-seven adult carriers of the m.3243A>G mutation were included in this study. The Newcastle Mitochondrial Disease Adult Scale was used for rating mitochondrial disease severity. In parallel, blood was drawn for GDF15 analysis by ELISA. Forty-nine carriers were included in a follow-up study. In a small subset of subjects of whom an echocardiogram was available from general patient care, myocardial deformation was assessed using two-dimensional speckle-tracking strain analysis.

A moderate positive correlation was found between the concentration of GDF15 and disease severity ( $r = 0.59$ ;  $p < 0.001$ ). The concentration of serum GDF15 was higher in m.3243A>G carriers with diabetes mellitus, cardiomyopathy and renal abnormalities. After a two-year follow-up, no significant correlation was found between the change in disease severity and the change in the concentration of GDF15 or between the GDF15 level at the first assessment and the change in disease severity. In the subcohort of patients of whom an echocardiogram was available, the concentration of GDF15 correlated moderately to longitudinal global strain ( $r = 0.55$ ;  $p = 0.006$ ;  $n = 23$ ) but not to circumferential or radial strain. Our results indicate that serum GDF15 is not a strong surrogate marker for general mitochondrial disease severity. Its value in indicating myocardial deformation should be confirmed in a prospective longitudinal study.

## Introduction

One of the key aspects of improving the quality of clinical trials is the identification of biomarkers that are indicative of clinically relevant outcome.<sup>1</sup> The perfect biomarker, correlating closely to clinical disease severity, would make the follow-up of patients easier, cheaper and less invasive, both in clinical trials and in regular patient care.<sup>2</sup> Moreover, since functional measures are subject to bias (including patient factors influencing performance, report bias, and inter- and intra-rater variability), measuring more objective disturbances of physiology may seem more reliable. Several tests have been used to indirectly measure the disturbed mitochondrial energy metabolism in patients with mitochondrial disease, including the determination of lactic acid concentration in the brain by magnetic resonance spectroscopy or serum and serum Fibroblast Growth Factor 21 (FGF21).<sup>3</sup> Although both lactic acid and FGF21 seemed to correlate to disease severity, the concentration during follow-up did not correlate to disease progression.<sup>4,5</sup>

A recent study reported growth and differentiation factor 15 (GDF15) as a potential new diagnostic biomarker for mitochondrial disease.<sup>6</sup> GDF15 was already known as a quite non-specific biomarker for cancer, as well as cardiac, pulmonary, renal and gynaecological disease.<sup>7-10,11,12</sup> However, the concentrations reported in these disorders are within the 1,000 – 7,000 pg/mL range,<sup>8,9,11,13,14</sup> whereas concentrations as high as 85,252 pg/mL were reported in patients with mitochondrial disease.<sup>6</sup> A child with the m.3243A>G mutation, the most commonly observed mutation leading to mitochondrial disease, was reported to have a concentration of 6,999 pg/mL (reference value 380 pg/mL (95% CI 59 - 701 pg/mL)).

To evaluate the value of GDF15 as a surrogate marker for disease severity and disease progression, we examined GDF15 in a large cohort of adult carriers of the m.3243A>G mutation. Since GDF15 was previously reported as a biomarker for symptoms associated with the m.3243A>G mutation, such as cardiomyopathy,<sup>15,16</sup> diabetes mellitus<sup>13</sup> and renal failure<sup>14,17,18</sup> we assessed these symptoms and organ functions in more detail.

## Methods

### Patients

We determined the serum GDF15 concentration in adult carriers of the m.3243A>G mutation included in our 'National inventory of patients with the m.3243A>G mutation' study. In each subject, the heteroplasmy percentage in buccal mucosa cells, urinary epithelial cells (UEC) and leukocytes was determined using pyrosequencing.<sup>19</sup> A heteroplasmy percentage  $\geq 5\%$  can be detected using this technique. Subjects with a detectable heteroplasmy percentage in either buccal mucosa cells, leukocytes, or UEC were considered to be carriers of the mutation. In this national inventory, clinical disease severity is monitored approximately two-yearly in both symptomatic and asymptomatic subjects carrying the m.3243A>G mutation.<sup>19</sup> Clinical disease severity is rated using the Newcastle Mitochondrial Disease Adult Scale (NMDAS), a multi-dimensional clinical scale encompassing current function (patient's opinion), system specific involvement (assessment of multi-system disease), and current clinical assessment (physical examination).<sup>20</sup> Carriers were rated as having asymptomatic (NMDAS = 0); mild (NMDAS 1-5), moderate (NMDAS 6-20) or severe (NMDAS >20) mitochondrial disease (cut-off values based on expert opinion). Seventy-six carriers were included in the follow-up study, serum of 50 of these carriers was available for analysis (Figure 1). For a more detailed description of the methods, we refer to our previous study on FGF21 concentrations in this population.<sup>4</sup> Patients with cancer and pregnant women were excluded since GDF15 is a known biomarker for these conditions. Since cardiomyopathy, diabetes mellitus and renal failure – for which GDF15 is also a biomarker – are highly prevalent in carriers of the m.3243A>G mutation,<sup>19</sup> we also evaluated the influence of these conditions on the GDF15 concentration. Microalbuminuria was defined as an albumin-to-creatinine-ratio of  $>2.0$  g/mol for men and  $>2.5$  g/mol for women, measured in a spot sample of urine. Decreased creatinine clearance was defined as a glomerular filtration rate  $<60$  ml/min/1.73m<sup>2</sup>. Carriers were classified as having decreased creatinine clearance only, microalbuminuria only, both or neither. The presence and severity of diabetes mellitus (DM) follows from the NMDAS (the presence of DM was rated as DM requiring diet or medication). The measurement of myocardial strain is explained in more detail later in this section.

Thirty non-carrier family members were included as a nuclear genetic and environmental matched reference population. The maternal relatives who showed no signs of diabetes mellitus, renal disease or cardiac disease were included in the study. In these subjects, heteroplasmy percentages  $\leq 4\%$  (the assay's detection limit) in UEC, leukocytes and/or buccal mucosa cells were established.

## GDF15

All samples were measured in duplicate, following the instructions of the manufacturer (*R&D Biosystems, Minneapolis, US*). The inter-assay and intra-assay variability for high and low values was determined based on high and low controls samples, respectively. The functional sensitivity determined at 20% covariance (CV) was 9.8 pg/ml. At a level of 1,695 pg/ml the within-assay CV was 3.5% and the between assay CV was 5.7%. At a level of 729 pg/ml the within-assay CV was 3.6% and the between-assay CV was 2.2%. Samples with an initial intra-assay covariance (CV) >15% were repeated. The two values obtained from the duplicated measurements were averaged, and the mean of these two measurements was used for further analysis. Samples with a concentration higher than the highest standard value were diluted and analysis was repeated. According to the kit's manufacturer, the assay has no cross reactivity with human GDF9 and GDF11. Age- and gender based reference values for serum GDF15 were adopted from the Framingham Offspring Study<sup>21</sup> (elevated GDF15 concentration is above the 97.5th percentile matched for age and gender).

## Medical ethical approval

This study ('National Inventory of Patients with the m.3243A>G mutation') was approved by the regional Medical Research Ethics Committee. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2013.<sup>22</sup> Informed consent was obtained from all patients before being included in the study.

## Myocardial strain measurement

As part of general patient care, carriers of the m.3243A>G mutation regularly undergo diagnostic echocardiography. Myocardial deformation was only assessed only if an echocardiogram was performed less than one year from an available GDF15 sampling. The myocardial strain, a measure for the deformation of the myocardium throughout the cardiac cycle, was determined using two-dimensional speckle-tracking strain analysis in accordance with a previously published protocol.<sup>23</sup> Since echocardiography was done as part of general patient care and 2D strain measurement can only be performed when the images are obtained following a specific protocol, only a subset of carriers was included in this part of the study. Strain values are dimensionless and are expressed in percentages. Global longitudinal left ventricular myocardial strain was calculated by averaging the six segments of the 4-chamber long-axis view. Global radial and circumferential myocardial strain were calculated by averaging the six segments of the mid-cavity short-axis view (at the level of the papillary muscles). If less than four out of six segments showed valid tracking of the myocardium (e.g. because of regional inferior image quality or poor tracking by the software), the strain measurement was excluded from our analyses.

## Statistics

The absolute difference between two parameters was calculated by subtracting the first measurement from the second measurement. All parameters were assessed for (log)normality. To prevent non-real values for zero values of the NMDAS including its subsections and symptom specific items, all values were increased by 1 prior to log transformation. The changes in GDF15 concentration and NMDAS score were increased by 3,000 and 10, respectively. Variables with a (log)normal distribution, were compared using parametric tests, the mean and 95%-confidence intervals are reported. Variables that deviated strongly from a (log)normal distribution were analysed by performing a non-parametric test, the median and interquartile ranges are reported. Correlation coefficients were interpreted in accordance with the guidelines provided at the BMJ website (<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>; consulted 31-Jul-2014). Thus, a correlation coefficient ( $r$ ) of 0.80 to 1.0 is considered a very strong relationship;  $r = 0.60$  to 0.79 is considered a strong relationship;  $r = 0.40$  to 0.59 is considered a moderate relationship;  $r = 0.20$  to 0.39 is considered a weak relationship; and  $r = 0.00$  to 0.19 is considered a very weak or no relationship.

Several covariates for GDF15 are known from literature,<sup>21</sup> including age, the presence of DM, smoking (covariates with higher estimated coefficient than 0.1) and renal failure. These, together with other possible clinical including covariates (gender, body mass index (BMI), heteroplasmy percentage in UEC and leukocytes and disease severity (NMDAS)) were included as candidate predictors for GDF15. Gender, age, BMI, heteroplasmy percentage in UEC and leukocytes and the concentration of GDF15 and FGF21 were considered candidate predictors for disease severity (NMDAS score). The influence of nominal and ordinal candidate predictors was determined by comparing between groups; the influence of continuous data was evaluated in a by linear regression analysis. The influence of cardiomyopathy was studied in more detail in a subgroup of carriers. For the correlation between the concentration of GDF15 and strain measurements, only echocardiography examinations performed no more than 1 year from sampling, were analysed. Forward and backward iterative multi-variate linear regression models were used to determine the influence of covariates and to determine the contribution of GDF15 in predicting or monitoring the disease course. Possible candidate predictors were only included for iterative multi-variate modelling if they correlated to the dependent variable during univariate correlation analysis ( $p < 0.1$ ). Standardised regression coefficients ( $\beta$ ) are presented for each variable.

Because 38 percent of the variance in the concentration of GDF15 is genetically determined,<sup>21</sup> we performed two additional analyses to correct for the effect of kinship in our analyses. First, we performed a separate analysis that included only the most severely affected patient in each family (in case two family members had the same



NMDAS score, we included the youngest person with that score, assuming a relatively more severe disease in this person as age related complaints are also included in the NMDAS). Secondly, we used generalised estimating equation models (working correlation structure: independence, with robust standard error for correction) to confirm the contribution of covariates and candidate predictors found by linear regression models, corrected for kinship.

Because genetic factors between family members are not likely to influence (intra-individual) changes in GDF15 levels during follow-up, we only used linear regression models to determine the influence of covariates on the change in disease severity and the concentration of GDF15 (longitudinal study). The (absolute) change in disease severity (*i.e.* NMDAS score) was used as a possible candidate predictor for the change in GDF15 concentration. Both the change in the concentration of GDF15 and the change in the concentration of FGF21 were included as candidate predictors for the change in disease severity (*i.e.* NMDAS score).

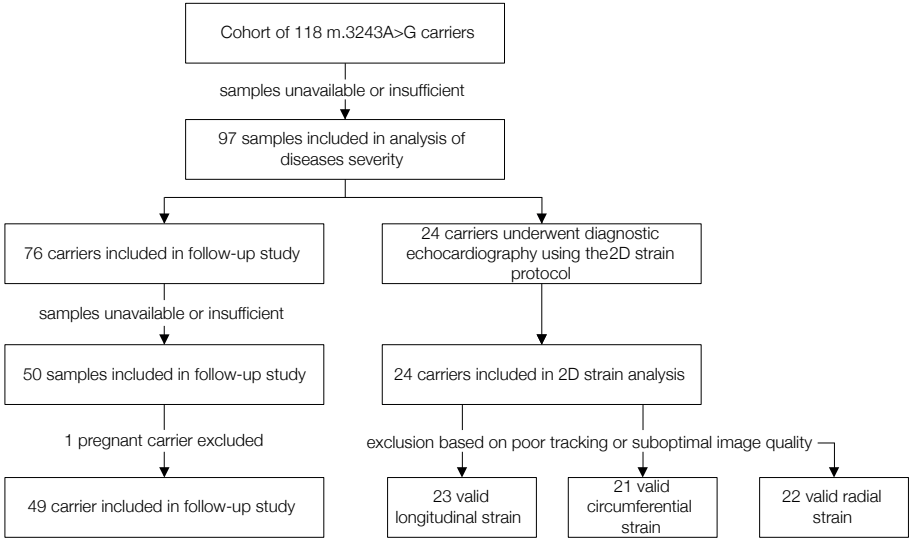
In case a high number of tests were performed (5 or more), critical  $p$ -values were adjusted using the Bonferroni method (*i.e.* critical  $p = 0.05/n$  where  $n$  = number of tests). All analyses were performed using IBM's SPSS statistics software packages, version 20.0.0.1.

## Results

### Patient characteristics

We initially included 118 adult subjects in our national inventory study. For a variety of reasons (*e.g.*, the samples were not available or had too small remaining volume for measuring GDF15), we were unable to determine the GDF15 levels in 21 of these subjects. No data were excluded because of high intra-assay covariance of GDF15 assessment. Thus, our final cohort included 97 adult carriers of the m.3243A>G mutation (Figure 1). The m.3243A>G mutation is not associated with a higher prevalence of cancer, and none of our subjects was known to suffer from any form of cancer at the time of the study. One patient did have a history of acute myeloid leukaemia (in remission for 8 years before samples were taken; [GDF15] 731 (baseline) and 1,926 (follow-up) pg/ml). One pregnant woman was excluded ([GDF15] 43,304 pg/ml) from the follow-up cohort.<sup>24</sup> For patient characteristics, we refer to Table 1. Renal function at the time of the sampling (spread: 6 months) was known in 86 patients (89% of total). Seventy-one percent of the total cohort had normal renal function, 17% had micro-albuminuria only, 2% had decreased creatinine clearance only and 9% suffered from both. Three patients had had a renal transplant, two of which had moderate transplant function (GFR 24 and 50 ml/min/1.73m<sup>2</sup>) and one had normal transplant function

(GFR >75 ml/min/1.73m<sup>2</sup>). The m.3243A>G carriers came from 41 distinct families (median: 2 subjects per family; range: 1-10 subjects per family). This cohort of adult carriers contained two asymptomatic patients (2%), 18 patients with mild mitochondrial disease (19%), 46 patients with moderate mitochondrial disease (47%), and 31 patients with severe mitochondrial disease (32%). Because sufficient material was not always available, heteroplasmy percentage in UEC is absent in two subjects; heteroplasmy percentage in leukocytes is absent for another two patients. In these four subjects, heteroplasmy levels ≥5% were established in the other available tissue (buccal mucosa cells, leucocytes or UEC).



**Figure 1.**  
**Flowchart of the study cohort**

**Table 1.**  
**Characteristics of the cohort of all adult m.3243A>G carriers at baseline, the follow-up cohort and the family matched controls.** Characteristics for all adult carriers ( $n = 97$ ) at baseline, for the patients within the follow-up cohort ( $n = 49$ ; A) at follow-up and for the family matched controls ( $n = 30$ ; B).  $p$ -values for the difference at baseline between the follow-up cohort and all carriers were calculated (difference at baseline; Chi square test (2-sided); Fisher's exact test (2-sided); Mann Whitney U test; independent sample  $t$ -test). Also,  $p$ -values for the difference between the first and the second measurement within the follow-up cohort are shown (difference from baseline; Wilcoxon signed rank test, paired  $t$ -test). The presence of Diabetes Mellitus was obtained from the NMDAS scale (score on Diabetes Mellitus item  $\geq 3$ ).

All carriers						Follow-up (49 carriers)						Difference at baseline	Difference from baseline
		central value	spread	range	$n =$	central value	spread	range	$n =$	$p =$	$p =$		
Gender	% female	71			97	65				0.20			
Age (years)	mean, 95%CI	45	17 - 72	18 - 81	97	45#	18 - 64	22 - 64	49	0.65			
BMI (kg/m <sup>2</sup> )	median, IQR	22.8	20.7 - 26.0	16.7 - 40.9	97	23.6#	17.6 - 33.2	17.6 - 33.1	49	0.96			
Smoking (current)	% yes	22			84	25			48	1.00			
Diabetes Mellitus (prevalence)	% yes	39			97	37#			49	0.62			
Heteroplasmy percentage leucocytes (%)	median, IQR	17	8 - 26	2 - 49	95	19#§	4 - 41	2 - 49	47	0.56			
Heteroplasmy percentage UEC (%)	median, IQR	45	27 - 69	4 - 96	95	52#§	7 - 91	5 - 96	47	0.1			
NMDAS score	mean, 95%CI	12*	1 - 67	0 - 56	97	14*	2 - 47	1 - 92	49	0.81	0.15		
Domain 1	median, IQR	4	1 - 9	0 - 23	97	6*\$	1 - 21	0 - 38	49	0.85	0.001		
Domain 2	median, IQR	5	3 - 9	0 - 22	97	5*\$	0 - 16	0 - 31	49	0.81	0.48		
Domain 3	median, IQR	3	1 - 6	0 - 20	97	3*\$	0 - 14	0 - 23	49	0.87	0.96		
Myopathy score	median, IQR	2	1 - 4	0 - 11	97	2*\$	0 - 12	0 - 14	49	0.96	0.11		
Encephalopathy score	median, IQR	2	0 - 3	0 - 12	97	1*\$	0 - 5	0 - 17	49	0.98	0.38		
Diabetes Mellitus (severity)	median, IQR	0	0 - 5	0 - 5	97	1*\$	0 - 3	0 - 5	49	0.37	0.002		
Cardiomyopathy (severity)	median, IQR	0	0 - 1	0 - 5	97	0	0 - 5	0 - 5	49	0.59	0.001		
Stroke-like episodes (severity)	median, IQR	0	0 - 0	0 - 5	97	0	0 - 0	0 - 5	49	0.98	0.58		
QoL mental	median, IQR	48	41 - 55	25 - 66	97	45	39 - 55	14 - 62	47	0.11	0.63		
QoL physical	median, IQR	41	33 - 51	17 - 61	97	41	35 - 49	21 - 64	47	0.29	0.36		
[GDF15] (pg/ml)	mean, 95%CI	1525*	411 - 5691	333-7421	97	1484#	1072 - 2577	370 - 7359	49	0.66	0.22		
[FGF21] (pg/ml)	median, IQR	263	142 - 534	3 - 1491	93	278	117 - 473	56 - 1776	24	0.96	0.27		

Family matched controls (n = 30)						Difference from all carriers
		central value	spread	range	n =	p =
Gender	% female	73			30	1.00
Age (years)	mean, 95%CI	41	20 - 68	18 - 68	30	0.518
BMI (kg/m <sup>2</sup> )	median, IQR	24.0*	19.7 - 33.0	19 - 37	30	<b>0.004</b>
Smoking (current)	% yes	8			13	0.286
Diabetes Mellitus (prevalence)	% yes	7			30	<b>0.001</b>
Heteroplasmy percentage leucocytes (%)	median, IQR	1§	0 - 3	0 - 3	30	<b>&lt;0.001</b>
Heteroplasmy percentage UEC (%)	median, IQR	1§	0 - 2	0 - 2	29	<b>&lt;0.001</b>
NMDAS score	mean, 95%CI	3*	0 - 17	0 - 22	30	<b>&lt;0.001</b>
Domain 1	median, IQR					
Domain 2	median, IQR					
Domain 3	mean, 95%CI					
Myopathy score	median, IQR					
Encephalopathy score						
Diabetes Mellitus (severity)						
Cardiomyopathy (severity)						
Stroke-like episodes (severity)						
QoL mental		49§	33 - 64	31 - 65	30	0.314
QoL physical		51	43 - 56	33 - 64	1	<b>&lt;0.001</b>
[GDF15] (pg/ml)		490*	272 - 1616	236 - 1687	30	<b>&lt;0.001</b>
[FGF21] (pg/ml)		13§	0 - 444	3 - 1072	25	<b>&lt;0.001</b>

BMI = body mass index; 95%CI = 95% confidence interval; Section 1 = current function; Section 2 = system specific involvement; Section 3 = current clinical assessment; Encephalopathy score = sum of the encephalopathic symptoms of the NMDAS (psychiatric symptoms, migraine, seizure, stroke-like episodes, encephalopathy and cognition); GDF15 = Growth and Differentiation Factor 15; IQR = interquartile range; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); n = number of carriers of which data were available at that specific time point; NMDAS = Newcastle Mitochondrial Disease Adult Scale; UEC = urinary epithelial cells; \*lognormal distribution; # at baseline; ¶ median and IQR instead of mean and 95%CI are given; § mean and 95%CI instead of median and IQR are given.

### The value of GDF15 as an indicator of clinical disease severity

Fifty-one carriers (53%) had an elevated concentration of GDF15 (i.e. higher than the 97.5<sup>th</sup> percentile of healthy age and sex matched controls) compared to the age- and gender matched reference population from literature.<sup>21</sup> Carriers with an elevated concentration of GDF15 had higher NMDAS scores compared to carriers with normal concentrations of GDF15 (2,440 pg/ml (95%CI 1,097 – 5,431 pg/ml) versus 902 pg/ml (95%CI 381 – 2,131 pg/ml);  $p < 0.001$ ). The correlation between the concentration of GDF15 and total NMDAS score in the cohort of m.3243A>G carriers at baseline was  $r = 0.59$  ( $p < 0.001$ ;  $n = 97$ ; Figure 2). This correlation coefficient is not significantly higher ( $p = 0.20$ ) than the correlation between NMDAS score and FGF21 ( $r = 0.45$ ;  $p < 0.001$ ;  $n = 93$ ) in this cohort) 95 out of 99 patients (96%) of the current cohort are the same as in the study of FGF21.<sup>4</sup> No significant correlation was found between the heteroplasmy percentage in UEC and the NMDAS score, nor between the heteroplasmy percentage in leukocytes and the NMDAS score. Among the 66 patients with asymptomatic, mild or moderate disease, the correlation coefficient between the total NMDAS score and GDF15 was 0.54 ( $p < 0.001$ ); among the 31 patients with severe disease, we found no significant correlation between the total NMDAS score and the concentration of GDF15. The correlation between the heteroplasmy level in UEC and GDF15 was 0.30 ( $p = 0.003$ ;  $n = 95$ ). The correlation between the heteroplasmy level in leukocytes and GDF15 was 0.22 ( $p = 0.031$ ;  $n = 95$ ).

Since the 97 adult carriers came from 41 families, we minimised the role of kinship in these associations by analysing only the most severely affected patient in each of the 41 families. The correlation coefficient between the total NMDAS score and GDF15 concentration in this cohort was  $r = 0.46$  ( $p = 0.003$ ) and was similar ( $p = 0.58$ ) compared to the full cohort. During iterative multi-variate linear modelling, the following parameters were found to be independent predictors for disease severity: concentration of GDF15, age, concentration of FGF21, and the heteroplasmy percentage in UEC ( $\beta(\text{GDF15}) = 0.38$  ( $p < 0.001$ );  $\beta(\text{age}) = 0.32$  ( $p < 0.001$ );  $\beta(\text{heteroplasmy UEC}) = 0.24$  ( $p = 0.005$ );  $\beta(\text{FGF21}) = 0.21$ ;  $p = 0.033$ ). When GDF15 and FGF21 were included as the only independent predictors for disease severity in a linear regression model, only GDF15 was included ( $\beta(\text{GDF15}) = 0.59$ ;  $p < 0.001$ ).

We found a moderate correlation between the concentration of GDF15 and the concentration of FGF21 ( $r = 0.54$ ;  $p > 0.001$ ;  $n = 93$ ).

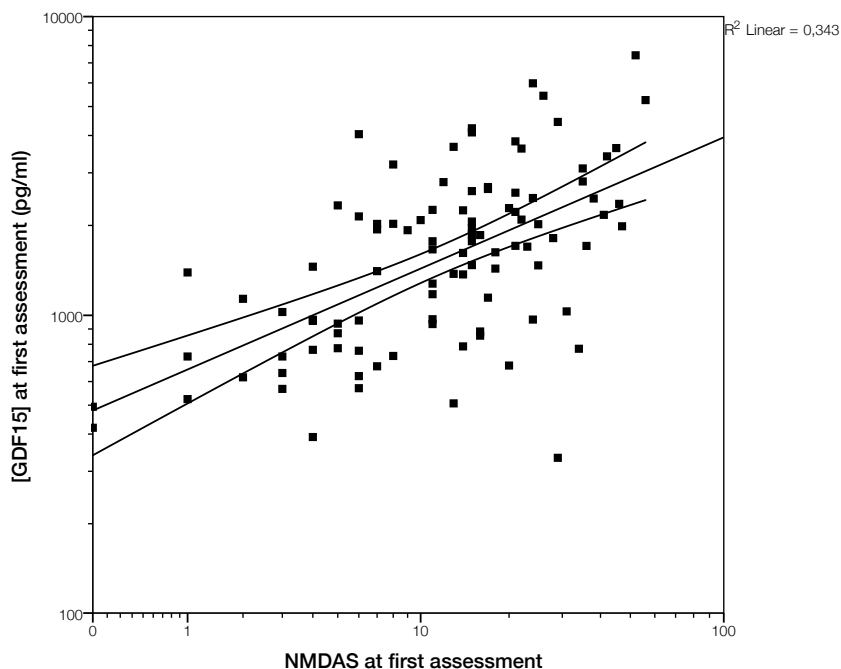


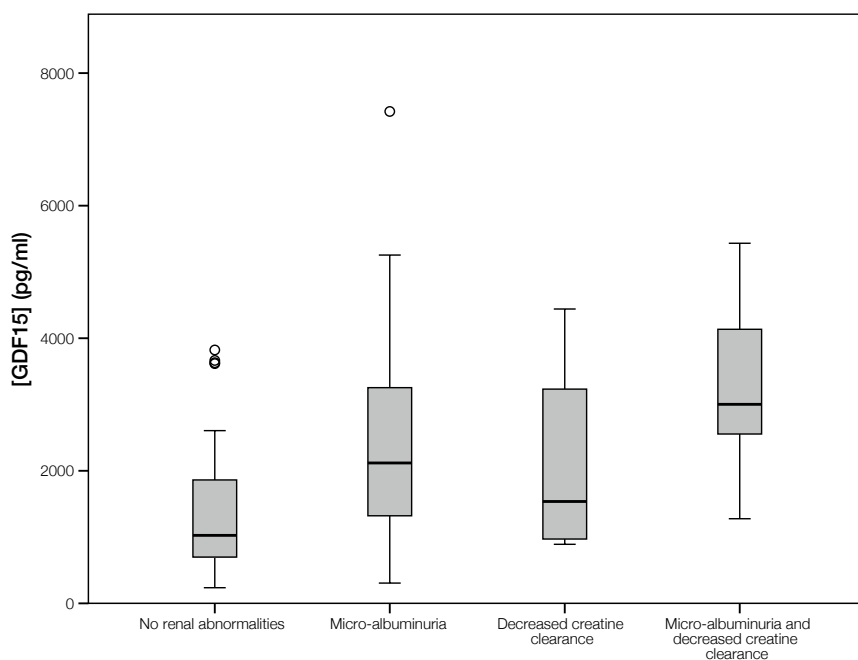
Figure 2.

**Correlation between the concentration of GDF15 and disease severity in adult m.3243A>G carriers.** The correlation between the concentration of GDF15 and the total NMDAS score (disease severity) is  $r = 0.59$  ( $p < 0.001$ ). Scales are loglinear. [GDF15] = the concentration of Growth and Differentiation Factor 15; NMDAS = Newcastle Mitochondrial Disease Scale

### Covariates for the concentration of GDF15

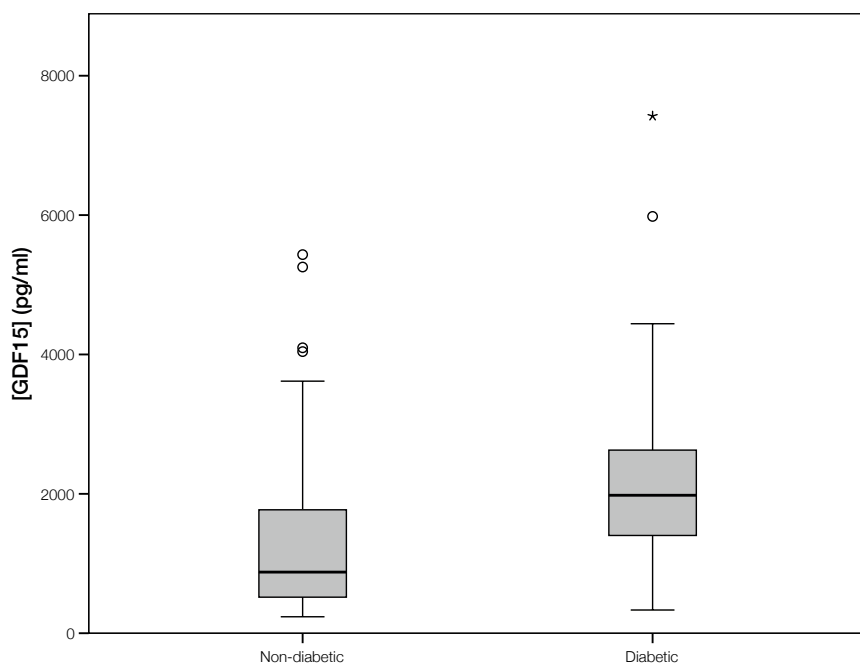
The concentration of GDF15 was not higher in females compared to males ( $p = 0.38$ ). The concentration of GDF15 was not higher among smokers compared to non-smokers ( $p = 0.70$ ;  $n = 84$ ). Patients without micro-albuminuria or decreased creatinine clearance had lower concentrations of GDF15 compared to patients with micro-albuminuria only or both micro-albuminuria or decreased creatinine clearance ( $p = 0.031$  and  $0.002$  respectively), but not compared to patients with decreased creatinine clearance only ( $p = 0.24$ ). The subgroups with renal abnormalities were comparable with respect to GDF15 concentrations (Figure 3). The concentration of GDF15 was significantly higher in patients with any kind of renal abnormalities compared to those without renal abnormalities (2,515 pg/ml (95%CI 699 - 9,045 pg/ml) versus 1,261 pg/ml (95%CI 373 - 4,264 pg/ml;  $n = 61$ );  $p < 0.001$ ). Carriers with micro-albuminuria had higher concentrations of GDF15 compared to those without micro-albuminuria (2,574 pg/ml (IQR 1,937 - 4,069 pg/ml;  $n = 25$ ) versus 1,406 pg/ml (IQR 781 - 1,973 pg/ml);  $p < 0.001$ ). Carriers with decreased creatinine clearance had higher concentrations of GDF15

compared to carriers with normal creatinine clearance including those with micro-albuminuria only (3,418 pg/ml (IQR 2,463 – 4,388 pg/ml;  $n = 12$ ) versus 1,471 pg/ml (IQR 867 – 2,105 pg/ml);  $p < 0.001$ ).



**Figure 3.**  
GDF15 concentrations in patients with and without renal abnormalities, including decreased creatinine clearance and micro-albuminuria.

Carriers with DM had higher concentrations of GDF15 compared to carriers without DM (1,958 pg/ml (95%CI 579 - 6,620 pg/ml) versus 1,299 pg/ml (95%CI 345 - 4,902 pg/ml);  $p = 0.003$ ; Figure 4). We found significantly higher serum GDF15 concentrations in patients with cardiomyopathy according to the NMDAS (including patients with asymptomatic ECG changes; NMDAS-Cardiomyopathy  $\geq 1$ ) compared to carriers without cardiomyopathy (2,574 pg/ml (998-6,638) versus 1,371 pg/ml (381-4,937);  $p < 0.001$ ). Univariate regression analysis showed that the total NMDAS score was the only significant contributor to the concentration of GDF15 ( $\beta(\text{NMDAS}) = 0.59$ ;  $p < 0.001$ ). In our cohort, we found no significant contribution of age and BMI to the concentration of GDF15. Generalised estimating equations confirmed the predictive value of disease severity as a significant contributor to the concentration of GDF15, after correcting for kinship clustering ( $p < 0.001$ ).



**Figure 4.**  
GDF15 concentrations in patients with and without diabetes mellitus.

### The value of GDF15 in predicting clinical disease progression

Approximately two years after the initial assessment, the GDF15 concentration and disease severity (NMDAS score) were measured again in 76 carriers from the initial cohort of 97 carriers. Of these 76 carriers, 50 samples were available (Figure 1). One carrier included in the follow-up group was excluded because of pregnancy. The remaining 49 carriers in the follow-up study were similar compared to the total cohort with respect to (distribution of) gender, age, BMI, GDF15 concentration, the presence and severity of DM, cardiomyopathy, stroke-like episodes, myopathy and encephalopathy, heteroplasmy percentage in UEC, heteroplasmy percentage in leucocytes and NMDAS total and subsections scores. The characteristics of the entire cohort of m.3243A>G carriers at baseline and the follow-up cohort are summarised in Table 1. In this follow-up cohort, we found no significant correlation between the change in the NMDAS score (*i.e.* the change in NMDAS score between the first and second follow-up visit) and the change in the concentration of GDF15 (*i.e.* the change in GDF15 concentration between the first and the second follow-up visit) ( $r = 0.006$ ;  $p = 0.97$ ;  $n = 49$ ). Moreover, no significant correlation was found between GDF15 concentration at the first visit and change in the NMDAS (disease progression) during follow-up ( $r = -0.19$ ;



$p = 0.18$ ;  $n = 49$ ). Linear regression also revealed that the change in total disease severity did not contribute significantly to the change in the GDF15 concentration ( $p = 0.72$ ). A linear regression model revealed that the change in FGF21 concentration, but not the change in GDF15 concentration contributed significantly to the change in disease severity (*i.e.* NMDAS score;  $\beta(\text{FGF21}) = 0.45$  ( $p = 0.03$ ;  $n = 24$ )). There was no correlation between the change in the concentration of GDF15 and the change in the concentration of FGF21 ( $r = -0.047$ ;  $p = 0.85$ ;  $n = 18$ ).

## Myocardial strain

Twenty-four subjects underwent echocardiographies as part of clinical care within one year of GDF15 sampling. These subjects were similar compared to the whole group with respect to gender, age, BMI and total NMDAS score and cardiac and DM sub scores (Table 2). Qualitative descriptions of gross echocardiography findings include (not mutually exclusive): left ventricle (LV) hypertrophy ( $n = 8$ ), LV systolic dysfunction ( $n = 3$ ); LV diastolic dysfunction ( $n = 8$ ); mild aortic regurgitation ( $n = 2$ ) and mild mitral regurgitation ( $n = 1$ ). Ten carriers had normal gross echocardiography findings.

The correlation coefficient between global longitudinal strain and the concentration of GDF15 was 0.55 ( $p = 0.006$ ;  $n = 23$ ). There was no correlation between global circumferential or radial strain and the concentration of GDF15 ( $r = 0.17$ ;  $p = 0.47$ ;  $n = 21$  and  $r = 0.20$ ;  $p = 0.37$ ;  $n = 22$ , respectively).

## Family matched controls

The family-matched controls ( $n = 30$ ) were similar to their maternal relatives carrying the m.3243A>G mutation with respect to age ( $p = 0.52$ ), gender ( $p = 1.0$ ) and smoking ( $p = 0.29$ ), but not with respect to BMI ( $p = 0.004$ ) and the presence and severity of DM ( $p = 0.001$ ; Table 1). The mean GDF15 concentration in the family-matched control group was 490 pg/ml (95%CI 272 – 1,616 pg/ml; range 236 – 1,687 pg/ml), which was significantly lower than their relatives carrying the m.3243A>G mutation (1,525 pg/ml (95%CI 411 - 5,691 pg/ml) range 333 - 7,421 pg/ml;  $p < 0.001$ ). Only one maternal family member had an elevated concentration of GDF15 (1,560 pg/ml in a 18-year-old female) compared to age- and gender matched controls. None of the maternal relatives was pregnant or known to have cancer, renal dysfunction or cardiac problems.

**Tabel 2.**

**Patient characteristics of the cohort of all adult m.3243A>G carriers of which myocardial strain was reported.** Patient characteristics for those adult carriers ( $n = 24$ ) for whom myocardial strain was measured. The difference at baseline between this cohort and all carriers was calculated. The presence of Diabetes Mellitus was obtained from the NMDAS scale (score on Diabetes Mellitus item  $\geq 3$ ).

		Cardiac 2D strain cohort (24 carriers)			Difference
		central value (spread)	range	$n =$	$p =$
Gender	% female	75		24	0.80
Age (years)	mean, 95%CI	44 (25 - 63)	24 - 64	24	0.80
BMI (kg/m <sup>2</sup> )	median, IQR	22.3 (21 - 26)	18.3 - 35.8	24	0.91
Smoking (current)	% yes	22		23	0.78
Diabetes Mellitus (prevalence)	% yes	45		24	0.48
Heteroplasmy percentage leucocytes (%)	mean, 95%CI	18* (7 - 50)	7 - 49	23	0.22
Heteroplasmy percentage UEC (%)	mean, 95%CI	56 (24 - 93)	23 - 96	24	0.08
NMDAS score	mean, 95%CI	20 (1 - 54)	0 - 56	24	0.11
Section 1	mean, 95%CI	5* (0 - 23)	0 - 23	24	0.11
Section 2	mean, 95%CI	8 (0 - 23)	0 - 23	24	0.015
Section 3	mean, 95%CI	3* (0 - 16)	0 - 18	24	0.97
Myopathy score	mean, 95%CI	3 (0 - 7)	0 - 7	24	0.57
Encephalopathy score	median, IQR	2 (1 - 5)	0 - 11	24	0.25
Diabetes Mellitus (severity)	mean, 95%CI	2 (0 - 5)	0 - 5	24	0.31
Cardiomyopathy (severity)	median, IQR	0 (0 - 2)	0 - 5	24	0.06
Stroke-like episodes (severity)	median, IQR	0 (0 - 0)	0 - 5	24	0.055
QoL mental	mean, 95%CI	44 (27 - 61)	25 - 61	24	0.055
QoL physical	mean, 95%CI	38 (18 - 59)	17 - 60	24	0.081
[GDF15] (pg/ml)	mean, 95%CI	1,687* (483 - 5,791)	420 - 5,981	24	0.41
[FGF21] (pg/ml)	median, IQR	321 (252 - 768)	4 - 14,91	24	0.084

BMI = body mass index; 95%CI = 95% confidence interval; Section 1 = current function; Section 2 = system specific involvement; Section 3 = current clinical assessment; Encephalopathy score = sum of the encephalopathic symptoms of the NMDAS (psychiatric symptoms, migraine, seizure, stroke-like episodes, encephalopathy and cognition). GDF15 = Growth and Differentiation Factor 15; IQR = interquartile range; Myopathy score = sum of the myopathic symptoms of the NMDAS (exercise intolerance, respiratory muscle weakness, ptosis, external ophthalmoplegia and myopathy); NMDAS = Newcastle Mitochondrial Disease Adult Scale; UEC = urinary epithelial cells. \*lognormal distribution.

## Discussion

This study explored the value of serum GDF15 in indicating and monitoring mitochondrial disease severity and disease progression in adult m.3243A>G carriers. We found that the concentration of serum GDF15 correlates moderately to disease severity but does not correlate to disease progression in m.3243A>G carriers. Analysis of data obtained in general patient care indicated that GDF15 might be a surrogate biomarker of left ventricular myocardial strain.

So far, a few studies have found promising results regarding the diagnostic properties of GDF15 in patients with mitochondrial disease.<sup>6,25</sup> None of these studies focused on the value of GDF15 as a surrogate marker for predicting or monitoring disease progression. We found normal (age- and gender matched<sup>21</sup>) concentrations of GDF15 in 47% of our carriers, including both symptomatic and asymptomatic individuals and in 97% of our family-matched controls. The family-matched controls (bearing the same nuclear genetic and environmental background as the carriers, but without the m.3243A>G detectable in UEC, leucocytes or buccal mucosa cells) had significantly lower GDF15 concentrations compared to their maternal relatives carrying the m.3243A>G mutation.

GDF15 has also been reported as a non-specific biomarker for many diseases, including cancer, cardiac, pulmonary, renal and gynaecological disease.<sup>7-10</sup> Since several of these conditions are highly prevalent in m.3243A>G carriers,<sup>19,26</sup> including kidney failure,<sup>12</sup> diabetes mellitus<sup>13</sup> and cardiomyopathy,<sup>11</sup> the contribution of these conditions to the concentration of GDF15 was studied in more detail. We observed higher concentrations of GDF15 in carriers with renal abnormalities (mainly micro-albuminuria) and diabetes mellitus.<sup>27</sup>

The myocardial strain measured in echocardiograms<sup>29</sup> collected as part of general patient care with a maximum of one year apart from GDF15 sampling, indicate that GDF15 may be a promising surrogate marker for myocardial deformity in patients with the m.3243A>G mutation. The correlation between the concentration of GDF15 and longitudinal myocardial strain, but not between the concentration of GDF15 and circumferential or radial strain<sup>28,30,31</sup> might be explained by the observation that changes in longitudinal myocardial strain precede changes in ejection fraction and global longitudinal strain is now part of the recommendations in the assessment of cardiac function, e.g. in monitoring chemotherapy.<sup>15,32-35</sup> In summary, GDF15 is a non-specific marker that seems to be highly influenced by several other symptoms frequently seen in this patient group.

The previously proposed biomarker FGF21 had no additional value to GDF15 in predicting and monitoring disease severity in m. 3243A>G carriers.<sup>6,36</sup> It seems, neither of the two parameters are useful as a surrogate marker for disease severity and disease progression.<sup>43</sup>

Strengths of our study include the high number of carriers with different levels of the same mutation in their mitochondrial genome, resulting in a heterogeneous multi-system disease that we quantified systematically using a standardised and quantitative follow-up, enabling us to draw tentative conclusions regarding the value of GDF15 as a prognostic biomarker for monitoring disease progression. Several limitations to our study are worth mentioning. The period of follow-up is relatively short for this slowly progressive and often oscillating disease. The NMDAS used in the follow-up of disease severity is not very accurate and lacks sensitivity, making it unfit to establish subtle changes, in this aspect the study is underpowered. The quality of the serum may have been influenced by the storage and previous use (freeze-thawing) of the samples used in our study. Since the echocardiograms for 2D strain analysis were collected as part of clinical care, the level of standardisation was suboptimal and samples for GDF15 analysis were not taken at the same moment as the echocardiography. Therefore, the results of this pilot study, including the responsiveness of GDF15 as a biomarker for myocardial strain, need to be confirmed in a prospective study. Finally, since only carriers of the m.3243A>G were included in this study, one may not extrapolate these findings to other causes of mitochondrial failure

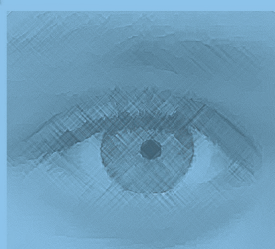
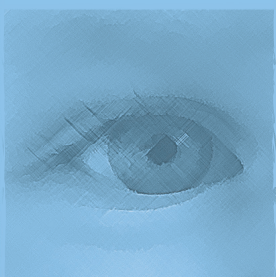
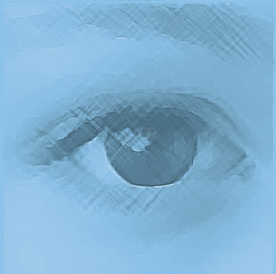
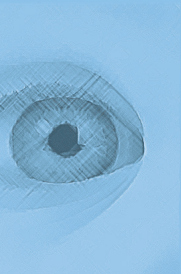
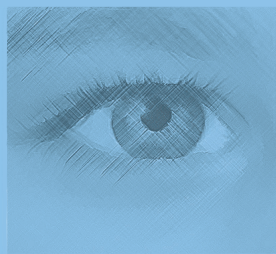
The mitochondrial disease field is diligently looking for easy-to-measure biomarkers, both for diagnostic and prognostic purposes. In this study, we have demonstrated that the concentration of serum GDF15 is moderately related to disease severity, but not to disease progression. Although the current study does not focus on the diagnostic applicability of GDF15, the lack of correlation between the concentration of GDF15 and disease progression in this two-year follow-up study makes this biomarker unsuitable as a prognostic biomarker. The moderate correlation between myocardial strain and the concentration of GDF15 is a promising starting point for finding a prognostic biomarker for myocardial deformity but warrants further confirmation, preferably in a longitudinal study including an in depth evaluation of renal function, glucose tolerance and the effects of pharmacological interventions.

## References

1. Pfeffer G, Horvath R, Klopstock T, *et al.* New treatments for mitochondrial disease-no time to drop our standards. *Nature reviews Neurology* 2013;9:474-81.
2. Mayeux R. Biomarkers: Potential Uses and Limitations. *NeuroRx* 2004;1:182-8.
3. Suomalainen A. Biomarkers for mitochondrial respiratory chain disorders. *Journal of inherited metabolic disease* 2011;34:277-82.
4. Koene S, de Laat P, van Tienoven DH, *et al.* Serum FGF21 levels in adult m.3243A>G carriers: Clinical implications. *Neurology* 2014 Jul 8;83(2):125-33.
5. Lee SK, Kim J, Kim HD, Lee JS, Lee YM. Initial experiences with proton MR spectroscopy in treatment monitoring of mitochondrial encephalopathy. *Yonsei Med J* 2010;51:672-5.
6. Kalko SG, Paco S, Jou C, *et al.* Transcriptomic profiling of TK2 deficient human skeletal muscle suggests a role for the p53 signalling pathway and identifies Growth and Differentiation Factor-15 as a potential novel biomarker for mitochondrial myopathies. *BMC genomics* 2014;15:91.
7. Yang CZ, Ma J, Zhu DW, *et al.* GDF15 is a potential predictive biomarker for TPF induction chemotherapy and promotes tumorigenesis and progression in oral squamous cell carcinoma. *Ann Oncol* 2014;24:24.
8. Trovik J, Salvesen HB, Cuppens T, Amant F, Staff AC. Growth differentiation factor-15 as biomarker in uterine sarcomas. *Int J Gynecol Cancer* 2014;24:252-9.
9. Izumiya Y, Hanatani S, Kimura Y, *et al.* Growth differentiation factor-15 is a useful prognostic marker in patients with heart failure with preserved ejection fraction. *Can J Cardiol* 2014;30:338-44.
10. Kempf T, Wollert KC. Risk stratification in critically ill patients: GDF15 scores in adult respiratory distress syndrome. *Crit Care* 2013;17:173.
11. Montoro-Garcia S, Hernandez-Romero D, Jover E, *et al.* Growth differentiation factor-15, a novel biomarker related with disease severity in patients with hypertrophic cardiomyopathy. *Eur J Intern Med* 2012;23:169-74.
12. Breit SN, Carrero JJ, Tsai VW, *et al.* Macrophage inhibitory cytokine-1 (MIC-1/GDF15) and mortality in end-stage renal disease. *Nephrol Dial Transplant* 2012;27:70-5.
13. Dominguez-Rodriguez A, Abreu-Gonzalez P, Avanzas P. Usefulness of Growth Differentiation Factor-15 Levels to Predict Diabetic Cardiomyopathy in Asymptomatic Patients With Type 2 Diabetes Mellitus. *Am J Cardiol* 2014;2:01367-8.
14. Ho JE, Hwang SJ, Wollert KC, *et al.* Biomarkers of cardiovascular stress and incident chronic kidney disease. *Clin Chem* 2013;59:1613-20.
15. Hollingsworth KG, Gorman GS, Trenell MI, *et al.* Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load. *Neuromuscul Disord* 2012;22:592-6.
16. Xu X, Li Z, Gao W. Growth differentiation factor 15 in cardiovascular diseases: from bench to bedside. *Biomarkers* 2011;16:466-75.
17. Emma F, Bertini E, Salvati L, Montini G. Renal involvement in mitochondrial cytopathies. *Paediatric nephrology (Berlin, Germany)* 2012;27:539-50.
18. Barak Y, Arnon S, Wolach B, *et al.* MELAS syndrome: peripheral neuropathy and cytochrome C-oxidase deficiency: a case report and review of the literature. *Isr J Med Sci* 1995;31:224-9.

19. de Laat P, Koene S, van den Heuvel LP, Rodenburg RJ, Janssen MC, Smeitink JA. Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A > G mutation. *J Inher Metab Dis*. 2012 Nov;35(6):1059-69.
20. Schaefer AM, Phoenix C, Elson JL, McFarland R, Chinnery PF, Turnbull DM. Mitochondrial disease in adults: a scale to monitor progression and treatment. *Neurology* 2006;66:1932-4.
21. Ho JE, Mahajan A, Chen MH, *et al*. Clinical and genetic correlates of growth differentiation factor 15 in the community. *Clin Chem* 2012;58:1582-91.
22. Association WM. Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. 64th WMA General Assembly, Fortaleza, Brazil 2013.
23. Bulten BF, Mavinkurve-Groothuis AM, de Geus-Oei LF, *et al*. Early myocardial deformation abnormalities in breast cancer survivors. *Breast Cancer Res Treat* 2014;146:127-35.
24. Moore AG, Brown DA, Fairlie WD, *et al*. The transforming growth factor- $\beta$  superfamily cytokine macrophage inhibitory cytokine-1 is present in high concentrations in the serum of pregnant women. *J Clin Endocrinol Metab* 2000;85:4781-8.
25. Yatsuga SK, Y. Growth differentiation factor 15 and Fibroblast Growth Factor 21: novel biomarkers for mitochondrial diseases. United Mitochondrial Disease Foundation Conference 2014.
26. Hirano M, Konishi K, Arata N, *et al*. Renal complications in a patient with A-to-G mutation of mitochondrial DNA at the 3243 position of leucine tRNA. *Intern Med* 2002;41:113-8.
27. Maassen JA. Mitochondrial diabetes: pathophysiology, clinical presentation, and genetic analysis. *Am J Med Genet* 2002;115:66-70.
28. Arslan D, Cihan T, Kose D, *et al*. Growth-differentiation factor-15 and tissue doppler imaging in detection of asymptomatic anthracycline cardiomyopathy in childhood cancer survivors. *Clin Biochem* 2013;46:1239-43.
29. Heger J, Schiegnitz E, von Waldthausen D, Anwar MM, Piper HM, Euler G. Growth differentiation factor 15 acts anti-apoptotic and pro-hypertrophic in adult cardiomyocytes. *J Cell Physiol* 2010;224:120-6.
30. Lok SI, Winkens B, Goldschmeding R, *et al*. Circulating growth differentiation factor-15 correlates with myocardial fibrosis in patients with non-ischaemic dilated cardiomyopathy and decreases rapidly after left ventricular assist device support. *Eur J Heart Fail* 2012;14:1249-56.
31. St John Sutton M, Ky B, Regner SR, *et al*. Longitudinal strain in Friedreich Ataxia: a potential marker for early left ventricular dysfunction. *Echocardiography (Mount Kisco, NY)* 2014;31:50-7.
32. Bates MG, Hollingsworth KG, Newman JH, *et al*. Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers. *Eur Heart J Cardiovasc Imaging* 2013;14:650-8.
33. Bates MG, Hollingsworth KG, Newman JH, *et al*. Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers. Defining cardiac adaptations and safety of endurance training in patients with m.3243A>G-related mitochondrial disease. *Eur Heart J Cardiovasc Imaging* 2013;14:650-8.
34. St John Sutton M, Ky B, Regner SR, *et al*. Longitudinal strain in Friedreich Ataxia: a potential marker for early left ventricular dysfunction. *Echocardiography* 2014;31:50-7.
35. Plana JC, Galderisi M, Barac A, *et al*. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* 2014;15:1063-93.

36. Zhang JD, Berntsen N, Roth A, Ebeling M. Data mining reveals a network of early-response genes as a consensus signature of drug-induced in vitro and in vivo toxicity. *Pharmacogenomics J* 2014;14:208-16.
37. Medina J, Yamada S, Kojima I. Identification of differentially expressed genes during proliferative response of the liver induced by follistatin. *Endocr J* 2009;56:1067-77.
38. Yang H, Choi HJ, Park SH, Kim JS, Moon Y. Macrophage inhibitory cytokine-1 (MIC-1) and subsequent urokinase-type plasminogen activator mediate cell death responses by ribotoxic anisomycin in HCT-116 colon cancer cells. *Biochem Pharmacol* 2009;78:1205-13.
39. Ribas F, Villarroya J, Hondares E, Giralt M, Villarroya F. FGF21 expression and release in muscle cells: involvement of MyoD and regulation by mitochondria-driven signalling. *The Biochemical journal* 2014;23:23.
40. Potthoff MJ, Inagaki T, Satapati S, *et al.* FGF21 induces PGC-1alpha and regulates carbohydrate and fatty acid metabolism during the adaptive starvation response. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106:10853-8.
41. Li PX, Wong J, Ayed A, *et al.* Placental transforming growth factor-beta is a downstream mediator of the growth arrest and apoptotic response of tumor cells to DNA damage and p53 over-expression. *J Biol Chem* 2000;275:20127-35.
42. Lawton LN, Bonaldo MF, Jelenc PC, *et al.* Identification of a novel member of the TGF-beta superfamily highly expressed in human placenta. *Gene* 1997;203:17-26.
43. Phoenix C, Schaefer AM, Elson JL, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord* 2006;16:814-20.





# Supplementary Documents



## Supplementary Document 1

## Jadad score for included studies.

Study	Intervention	Disease	n	Total Jadad score
Klopstock <i>et al.</i> (2011) <sup>1</sup>	Idebenone	LHON	85	5
Kaufmann <i>et al.</i> (2006) <sup>2</sup>	Dichloroacetate	MELAS (m.3243A>G)	30	5
Kornblum <i>et al.</i> (2005) <sup>3</sup>	Creatine	PEO and KSS (single deletion)	15	5
Liet <i>et al.</i> (2003) <sup>4</sup>	Dimethylglycine	SLSJ-COX deficiency	5	5
Stacpoole <i>et al.</i> (2006) <sup>5</sup>	Dichloroacetate	Congenital lactic acidosis	43	4
Mancuso <i>et al.</i> (2010) <sup>6</sup>	Glutamylcysteine	PEO	13	4
Glover <i>et al.</i> (2010) <sup>7</sup>	Coenzyme Q10	Various MM	30	3
Klopstock <i>et al.</i> (2000) <sup>8</sup>	Creatine	PEO and MM	16	3
De Stefano <i>et al.</i> (1995) <sup>9</sup>	Dichloroacetate	Various syndromes	11	3
Tarnopolsky <i>et al.</i> (1997) <sup>10</sup>	Creatine	MELAS or MM	7	3
Rodriguez <i>et al.</i> (2007) <sup>11</sup>	Creatine, coenzyme Q10, lipoic acid	Various syndromes	16	2
Bresolin <i>et al.</i> (1990) <sup>12</sup>	Coenzyme Q10	Various MM	16	2
Duncan <i>et al.</i> (2004) <sup>13</sup>	Dichloroacetate	Various syndromes	9	2
Chen <i>et al.</i> (1997) <sup>14</sup>	Coenzyme Q10	Various MM	8	2
Barshop <i>et al.</i> (2004) <sup>15</sup>	Dichloroacetate	Various syndromes	37	1
Cejudo <i>et al.</i> (2005) <sup>16</sup>	Exercise	PEO and MM	20	1
Martinelli <i>et al.</i> (2012) <sup>17</sup>	EPI-743	Leigh syndrome	10	1
Taivassalo <i>et al.</i> (2006) <sup>18</sup>	Exercise	PEO (single deletion)	8	1
Remes <i>et al.</i> (2002) <sup>19</sup>	Coenzyme Q10 and nicotinamide	MELAS (m.3243A>G)	7	1
Sadun <i>et al.</i> (2012) <sup>20</sup>	EPI-743	LHON	5	1
Komura <i>et al.</i> (2003) <sup>21</sup>	Creatine	Various MM	5	1
Suzuki <i>et al.</i> (1998) <sup>22</sup>	Coenzyme Q10	MIDD (m.3243A>G)	44	0
Stacpoole <i>et al.</i> (2008) <sup>23</sup>	Dichloroacetate	Congenital lactic acidosis	36	0
Mashima <i>et al.</i> (2000) <sup>24</sup>	Idebenone, vitamins B2 and C	LHON	28	0
Koga <i>et al.</i> (2005) <sup>25</sup>	L-arginine	MELAS (m.3243A>G)	24	0
Blankenberg <i>et al.</i> (2012) <sup>26</sup>	EPI-743	Various syndromes	22	0
Campos <i>et al.</i> (1993) <sup>27</sup>	Carnitine	Various MM	21	0
Mathews <i>et al.</i> (1993) <sup>28</sup>	Coenzyme Q10	Various MM	16	0
Panetta <i>et al.</i> (2004) <sup>29</sup>	Coenzyme Q10, thiamine, riboflavin, vitamin C, high fat diet	Various syndromes	15	0

Enns <i>et al.</i> (2012) <sup>30</sup>	EPI-743	Various syndromes	14	0
Gold <i>et al.</i> (1996) <sup>31</sup>	Coenzyme Q10	Various MM	8	0
Ogasahara <i>et al.</i> (1986) <sup>32</sup>	Coenzyme Q10	KSS	5	0
Bernsen <i>et al.</i> (1993) <sup>33</sup>	Riboflavin	Complex I deficiency	5	0
Mori <i>et al.</i> (2004) <sup>34</sup>	Dichloroacetate	MELAS (m.3243A>G)	4	0

## References

1. Klopstock T *et al.* A randomized placebo-controlled trial of idebenone in Leber's hereditary optic neuropathy. *Brain* 2011;134:2677–86.
2. Kaufmann P *et al.* Dichloroacetate causes toxic neuropathy in MELAS: a randomized, controlled clinical trial. *Neurology* 2006;66:324–30.
3. Kornblum C *et al.* Creatine has no beneficial effect on skeletal muscle energy metabolism in patients with single mitochondrial DNA deletions: a placebo-controlled, double-blind <sup>31</sup>P-MRS cross-over study. *Eur J Neurol* 2005;12:300–9.
4. Liet JM *et al.* The effect of short-term dimethylglycine treatment on oxygen consumption in cytochrome oxidase deficiency: a double-blind randomized cross-over clinical trial. *J Paediatr* 2003;142:62–6.
5. Stacpoole PW *et al.* Controlled clinical trial of dichloroacetate for treatment of congenital lactic acidosis in children. *Paediatrics* 2006;117:1519–31.
6. Mancuso M *et al.* Oxidative stress biomarkers in mitochondrial myopathies, basally and after cysteine donor supplementation. *J Neurol* 2010;257:774–81.
7. Glover EI *et al.* A randomized trial of coenzyme Q10 in mitochondrial disorders. *Muscle Nerve* 2010;42:739–48.
8. Klopstock T *et al.* A placebo-controlled cross-over trial of creatine in mitochondrial diseases. *Neurology* 2000;55:1748–51.
9. De Stefano N *et al.* Short-term dichloroacetate treatment improves indices of cerebral metabolism in patients with mitochondrial disorders. *Neurology* 1995;45:1193–8.
10. Tarnopolsky MA, Roy BD & MacDonald JR. Randomised control trial of creatine monohydrate in patients with mitochondrial cytopathies. *Muscle Nerve* 1997;20:1502–9.
11. Rodriguez MC *et al.* Beneficial effects of creatine, CoQ10, and lipoic acid in mitochondrial disorders. *Muscle Nerve* 2007;35:235–42.
12. Bresolin N *et al.* Ubidecarenone in the treatment of mitochondrial myopathies: a multi-centre double-blind trial. *J Neurol Sci* 1990;100:70–8.
13. Duncan GE, Perkins LA, Theriaque DW, Neiberger RE & Stacpoole PW. Dichloroacetate therapy attenuates the blood lactate response to submaximal exercise in patients with defects in mitochondrial energy metabolism. *J Clin Endocrinol Metab* 2004;89:1733–8.
14. Chen RS, Huang CC & Chu NS. Coenzyme Q10 treatment in mitochondrial encephalomyopathies. Short-term double-blind, cross-over study. *Eur Neurol* 1997;37:212–8.
15. Barshop BA *et al.* Chronic treatment of mitochondrial disease patients with dichloroacetate. *Mol Genet Metab* 2004;83:138–49.
16. Cejudo P *et al.* Exercise training in mitochondrial myopathy: a randomized controlled trial. *Muscle Nerve* 2005;32:342–50.
17. Martinelli D *et al.* EPI-743 reverses the progression of the paediatric mitochondrial disease--genetically defined Leigh Syndrome. *Mol Genet Metab* 2012;107:383–8.
18. Taivassalo T *et al.* Endurance training and detraining in mitochondrial myopathies due to single large-scale mtDNA deletions. *Brain* 2006;129:3391–401.
19. Remes AM *et al.* Ubiquinone and nicotinamide treatment of patients with the 3243A-->G mtDNA mutation. *Neurology* 2002;59:1275–7.

20. Sadun AA *et al.* Effect of EPI-743 on the clinical course of the mitochondrial disease Leber hereditary optic neuropathy. *Arch Neurol* 2012;69:331–8.
21. Komura K, Hobbiebrunken E, Wilichowski EK & Hanefeld FA. Effectiveness of creatine monohydrate in mitochondrial encephalomyopathies. *Paediatr Neurol* 2003;28:53–8.
22. Suzuki S *et al.* The effects of coenzyme Q10 treatment on maternally inherited diabetes mellitus and deafness, and mitochondrial DNA 3243 (A to G) mutation. *Diabetologia* 1998;41:584–8.
23. Stacpoole PW *et al.* Evaluation of long-term treatment of children with congenital lactic acidosis with dichloroacetate. *Paediatrics* 2008;121:e1223–8.
24. Mashima Y, Kigasawa K, Wakakura M & Oguchi Y. Do idebenone and vitamin therapy shorten the time to achieve visual recovery in Leber hereditary optic neuropathy? *J Neuroophthalmol* 2000;20:166–70.
25. Koga Y *et al.* L-arginine improves the symptoms of stroke-like episodes in MELAS. *Neurology* 2005;64:710–2.
26. Blankenberg FG *et al.* Brain uptake of Tc99m-HMPAO correlates with clinical response to the novel redox modulating agent EPI-743 in patients with mitochondrial disease. *Mol Genet Metab* 2012;107:690–9.
27. Campos Y *et al.* Plasma carnitine insufficiency and effectiveness of L-carnitine therapy in patients with mitochondrial myopathy. *Muscle Nerve* 1993;16:150–3.
28. Mathews PM, Andermann F, Silver K, Karpai G & Arnold DL. Proton MR spectroscopic characterization of differences in regional brain metabolic abnormalities in mitochondrial encephalomyopathies. *Neurology* 1993;43:2484–90.
29. Panetta J, Smith LJ & Boneh A. Effect of high-dose vitamins, coenzyme Q and high-fat diet in paediatric patients with mitochondrial diseases. *J Inher Metab Dis* 2004;27:487–98.
30. Enns GM *et al.* Initial experience in the treatment of inherited mitochondrial disease with EPI-743. *Mol Genet Metab* 2012;105:91–102.
31. Gold R *et al.* Phosphorus magnetic resonance spectroscopy in the evaluation of mitochondrial myopathies: results of a 6-month therapy study with coenzyme Q. *Eur Neurol* 1996;36:191–6.
32. Ogasahara S *et al.* Treatment of Kearns-Sayre syndrome with coenzyme Q10. *Neurology* 1986 Jan;36(1):45–53.
33. Berssen PL, Gabreels FJ, Ruitenbeek W & Hamburger HL. Treatment of complex I deficiency with riboflavin. *J Neurol Sci* 1993;118:181–7.
34. Mori M, Yamagata T, Goto T, Saito S & Momoi MY. Dichloroacetate treatment for mitochondrial cytopathy: long-term effects in MELAS. *Brain Dev* 2004;26:453–8.
35. Koga Y *et al.* Effects of L-arginine on the acute phase of strokes in three patients with MELAS. *Neurology* 2002;58:827–8.

## Supplementary Document 2

### Summary of *p* values for clinically relevant end point.

Study	Jadad score	Outcome	<i>p</i> value	Comment
Klopstock <i>et al.</i> (2011) <sup>1</sup>	5	Best visual acuity	0.29	–
Kaufmann <i>et al.</i> (2006) <sup>2</sup>	5	GATE score (various time points)	0.16–0.78	–
Kornblum <i>et al.</i> (2005) <sup>3</sup>	5	Modified Boston score	NS	No <i>p</i> value was reported
Liet <i>et al.</i> (2003) <sup>4</sup>	5	Oxygen consumption	0.08	"Trend" favouring placebo
Stacpoole <i>et al.</i> (2006) <sup>5</sup>	4	GATE score	0.91	–
Mancuso <i>et al.</i> (2010) <sup>6</sup>	4	MRC score	NS	No <i>p</i> value was reported
Mancuso <i>et al.</i> (2010) <sup>6</sup>	4	SF36 score	NS	No <i>p</i> value was reported
Glover <i>et al.</i> (2010) <sup>7</sup>	3	ADL score	0.26	–
Glover <i>et al.</i> (2010) <sup>7</sup>	3	QoL score	0.09	–
Klopstock <i>et al.</i> (2000) <sup>8</sup>	3	MRC scale	0.14	–
Klopstock <i>et al.</i> (2000) <sup>8</sup>	3	NSS score	0.42	–
Klopstock <i>et al.</i> (2000) <sup>8</sup>	3	Hammersmith scale	0.92	–
Rodriguez <i>et al.</i> (2007) <sup>9</sup>	3	N/A	–	No relevant clinical end point or any single primary end point
De Stefano <i>et al.</i> (1995) <sup>10</sup>	3	Neurological strength/ gait examination score	>0.1	–
Tarnopolsky <i>et al.</i> (1997) <sup>11</sup>	3	ADL score	NS	No <i>p</i> value was reported
Tarnopolsky <i>et al.</i> (1997) <sup>11</sup>	3	2-minute walk	NS	No <i>p</i> value was reported
Cejudo <i>et al.</i> (2005) <sup>12</sup>	2	Functional exercise capacity	<0.01	–
Cejudo <i>et al.</i> (2005) <sup>12</sup>	2	Nottingham Health Profile	NS	No <i>p</i> value was reported
Bresolin <i>et al.</i> (1990) <sup>13</sup> (double-blind part only)	2	Global MRC scale	NS	No <i>p</i> value was reported
Bresolin <i>et al.</i> (1990) <sup>13</sup> (open-label part only)	0	Global MRC scale	0.01	–
Duncan <i>et al.</i> (2004) <sup>14</sup>	2	N/A	–	No relevant clinical end point or any single primary end point
Chen <i>et al.</i> (1997) <sup>15</sup>	2	Global MRC scale	0.05	–

Chen <i>et al.</i> (1997) <sup>15</sup>	2	ADL score	'Trend' towards improvement	We assumed a <i>p</i> value of 0.1
Barshop <i>et al.</i> (2004) <sup>16</sup>	1	Structured neurological-examination score	0.066	–
Barshop <i>et al.</i> (2004) <sup>16</sup>	1	Structured neurological symptom score	0.003	–
Martinelli <i>et al.</i> (2012) <sup>17</sup>	1	NPMDs score	0.02	–
Martinelli <i>et al.</i> (2012) <sup>17</sup>	1	PedsQL score	0.006	–
Martinelli <i>et al.</i> (2012) <sup>17</sup>	1	GMFM score (various time points)	0.005–0.007	–
Taivassalo <i>et al.</i> (2006) <sup>18</sup>	1	SF36 score	0.05	–
Remes <i>et al.</i> (2002) <sup>19</sup>	1	CIBIC score	NS	No <i>p</i> value was reported
Sadun <i>et al.</i> (2012) <sup>20</sup>	1	N/A	–	Multiple visual parameters reported in individual patients; no <i>p</i> value was reported
Komura <i>et al.</i> (2003) <sup>21</sup>	1	N/A	–	Clinically relevant end points were reported (time to climb 13 steps or run 9 m) but raw data not given, and no <i>p</i> values reported
Suzuki <i>et al.</i> (1998) <sup>22</sup>	0	Prevention of hearing loss	0.02	–
Stacpoole <i>et al.</i> (2008) <sup>23</sup>	0	QoL score	NS	No <i>p</i> value was reported
Mashima <i>et al.</i> (2000) <sup>24</sup>	0	Visual recovery >0.3	NS	No <i>p</i> value was reported
Mashima <i>et al.</i> (2000) <sup>24</sup>	0	Time to visual recovery >0.3	0.01	–
Koga <i>et al.</i> (2005) <sup>25</sup>	0	Clinical disability score	0.05	–
Blankenburg <i>et al.</i> (2012) <sup>26</sup>	0	NPMDs score	0.13	<i>p</i> value was calculated ourselves using raw data given in the paper: two-sided grouped student t-test
Campos <i>et al.</i> (1993) <sup>27</sup>	0	N/A	–	Subjective assessments only; no <i>p</i> value was reported
Mathews <i>et al.</i> (1993) <sup>28</sup>	0	N/A	–	No relevant clinical end point or any single primary end point
Panetta <i>et al.</i> (2004) <sup>29</sup>	0	N/A	–	Subjective assessments only; no <i>p</i> value was reported



Enns <i>et al.</i> (2012) <sup>30</sup>	0	NPMDS score	0.004	–
Gold <i>et al.</i> (1996) <sup>31</sup>	0	N/A	–	No relevant clinical end point or any single primary end point
Ogasahara <i>et al.</i> (1986) <sup>32</sup>	0	N/A	–	No relevant clinical end point or any single primary end point
Bernsen <i>et al.</i> (1993) <sup>33</sup>	0	N/A	–	No relevant clinical end point or any single primary end point
Mori <i>et al.</i> (2004) <sup>34</sup>	0	N/A	–	No relevant clinical end point or any single primary end point
Koga <i>et al.</i> (2002) <sup>35</sup>	0	Clinical disability score	0.05	–

ADL, activities of daily living; CIBIC, Clinician Interview-Based Impression of Change; GATE, Global Assessment of Treatment Efficacy; GMFM, Gross Motor Function Measure; MRC, Medical Research Council; N/A, not applicable; NPMDS, Newcastle Paediatric Mitochondrial Disease Scale; NS, not significant; SF36; NSS, Neuromuscular Symptom Scale; PedsQL, Paediatric Quality of Life Inventory; QoL, quality of life. Short Form 36 Health Survey

## References

1. Klopstock T *et al.* A randomised placebo-controlled trial of idebenone in Leber's hereditary optic neuropathy. *Brain* 2011;134:2677–86.
2. Kaufmann P *et al.* Dichloroacetate causes toxic neuropathy in MELAS: a randomised, controlled clinical trial. *Neurology* 2006;66:324–30.
3. Kornblum C *et al.* Creatine has no beneficial effect on skeletal muscle energy metabolism in patients with single mitochondrial DNA deletions: a placebo-controlled, double-blind <sup>31</sup>P-MRS cross-over study. *Eur J Neurol* 2005;12:300–9.
4. Liet JM *et al.* The effect of short-term dimethylglycine treatment on oxygen consumption in cytochrome oxidase deficiency: a double-blind randomised cross-over clinical trial. *J Paediatr* 2003;142:62–6.
5. Stacpoole PW *et al.* Controlled clinical trial of dichloroacetate for treatment of congenital lactic acidosis in children. *Paediatrics* 2006;117:1519–31.
6. Mancuso M *et al.* Oxidative stress biomarkers in mitochondrial myopathies, basally and after cysteine donor supplementation. *J Neurol* 2010;257:774–81.
7. Glover EI *et al.* A randomised trial of coenzyme Q10 in mitochondrial disorders. *Muscle Nerve* 2010;42:739–48.
8. Klopstock T *et al.* A placebo-controlled cross-over trial of creatine in mitochondrial diseases. *Neurology* 2000;55:1748–51.
9. De Stefano N *et al.* Short-term dichloroacetate treatment improves indices of cerebral metabolism in patients with mitochondrial disorders. *Neurology* 1995;45:1193–8.
10. Tarnopolsky MA, Roy BD & MacDonald JR. Randomised control trial of creatine monohydrate in patients with mitochondrial cytopathies. *Muscle Nerve* 1997;20:1502–9.
11. Rodríguez MC *et al.* Beneficial effects of creatine, CoQ10, and lipoic acid in mitochondrial disorders. *Muscle Nerve* 2007;35:235–42.
12. Bresolin N *et al.* Ubidecarenone in the treatment of mitochondrial myopathies: a multi-centre double-blind trial. *J Neurol Sci* 1990;100:70–8.
13. Duncan GE, Perkins LA, Theriaque DW, Neiberger RE & Stacpoole PW. Dichloroacetate therapy attenuates the blood lactate response to submaximal exercise in patients with defects in mitochondrial energy metabolism. *J Clin Endocrinol Metab* 2004;89:1733–8.
14. Chen RS, Huang CC & Chu NS. Coenzyme Q10 treatment in mitochondrial encephalomyopathies. Short-term double-blind, cross-over study. *Eur Neurol* 1997;37:212–8.
15. Barshop BA *et al.* Chronic treatment of mitochondrial disease patients with dichloroacetate. *Mol Genet Metab* 2004;83:138–49.
16. Cejudo P *et al.* Exercise training in mitochondrial myopathy: a randomised controlled trial. *Muscle Nerve* 2005;32:342–50.
17. Martinelli D *et al.* EPI-743 reverses the progression of the paediatric mitochondrial disease--genetically defined Leigh Syndrome. *Mol Genet Metab* 2012;107:383–8.
18. Taivassalo T *et al.* Endurance training and detraining in mitochondrial myopathies due to single large-scale mtDNA deletions. *Brain* 2006;129:3391–401.
19. Remes AM *et al.* Ubiquinone and nicotinamide treatment of patients with the 3243A-->G mtDNA mutation. *Neurology* 2002;59:1275–7.

20. Sadun AA *et al.* Effect of EPI-743 on the clinical course of the mitochondrial disease Leber hereditary optic neuropathy. *Arch Neurol* 2012;69:331–8.
21. Komura K, Hobbiebrunken E, Wilichowski EK & Hanefeld FA. Effectiveness of creatine monohydrate in mitochondrial encephalomyopathies. *Paediatr Neurol* 2003;28:53–8.
22. Suzuki S *et al.* The effects of coenzyme Q10 treatment on maternally inherited diabetes mellitus and deafness, and mitochondrial DNA 3243 (A to G) mutation. *Diabetologia* 1998;41:584–8.
23. Stacpoole PW *et al.* Evaluation of long-term treatment of children with congenital lactic acidosis with dichloroacetate. *Paediatrics* 2008;121:e1223–8.
24. Mashima Y, Kigasawa K, Wakakura M & Oguchi Y. Do idebenone and vitamin therapy shorten the time to achieve visual recovery in Leber hereditary optic neuropathy? *J Neuroophthalmol* 2000;20:166–70.
25. Koga Y *et al.* L-arginine improves the symptoms of strokelike episodes in MELAS. *Neurology* 2005;64:710–2.
26. Blankenberg FG *et al.* Brain uptake of Tc99m-HMPAO correlates with clinical response to the novel redox modulating agent EPI-743 in patients with mitochondrial disease. *Mol Genet Metab* 2012;107:690–9.
27. Campos Y *et al.* Plasma carnitine insufficiency and effectiveness of L-carnitine therapy in patients with mitochondrial myopathy. *Muscle Nerve* 1993;16:150–3.
28. Mathews PM, Andermann F, Silver K, Karpati G & Arnold DL. Proton MR spectroscopic characterization of differences in regional brain metabolic abnormalities in mitochondrial encephalomyopathies. *Neurology* 1993;43:2484–90.
29. Panetta J, Smith LJ & Boneh A. Effect of high-dose vitamins, coenzyme Q and high-fat diet in paediatric patients with mitochondrial diseases. *J Inher Metab Dis* 2004;27:487–98.
30. Enns GM *et al.* Initial experience in the treatment of inherited mitochondrial disease with EPI-743. *Mol Genet Metab* 2012;105:91–102.
31. Gold R *et al.* Phosphorus magnetic resonance spectroscopy in the evaluation of mitochondrial myopathies: results of a 6-month therapy study with coenzyme Q. *Eur Neurol* 1996;36:191–6.
32. Ogasahara S *et al.* Treatment of Kearns-Sayre syndrome with coenzyme Q10. *Neurology* 1986 Jan;36(1):45–53.
33. Bernsen PL, Gabreels FJ, Ruitenbeek W & Hamburger HL. Treatment of complex I deficiency with riboflavin. *J Neurol Sci* 1993;118:181–7.
34. Mori M, Yamagata T, Goto T, Saito S & Momoi MY. Dichloroacetate treatment for mitochondrial cytopathy: long-term effects in MELAS. *Brain Dev* 2004;26:453–8.
35. Koga Y *et al.* Effects of L-arginine on the acute phase of strokes in three patients with MELAS. *Neurology* 2002;58:827–8.

## Supplementary Document 3

Questionnaires (English translation) assessing i) complaints and most burdensome complaints, ii) limitations in body functions and iii) limitations in activity and participation in Dutch.

### i) Complaints and symptoms present in children with mitochondrial disease.

Parents and their children were asked to answer the following questions (three separate questionnaires):

The complaints my child experiences include:

The **three** complaints (I expect my child) to find most burdensome and most wanted to change include:

The **three** complaints I find most burdensome and most wanted to change include:

- Tiredness
- Lack of energy
- Frequent infections
- Delayed development
- Diminished intelligence
- Growth failure
- Diminished interaction
- Muscle weakness
- Low muscle power
- Stiff arms and legs
- Coordination problems
- Involuntary movements
- Restlessness
- Balance problems
- Epilepsy
- Vision problems
- Hearing problems
- Speech problems
- Swallowing problems
- Breathing problems
- Muscle pain
- Headache
- Pain in stomach/belly
- Vomiting
- Diarrhoea
- Constipation

- Behavioural problems
- Concentration problems
- Depressed mood
- Other, namely.....
- Other, namely.....
- Other, namely.....

## ii ) and iii) Checklist International Classification of Functioning (under 3 years)

This is a questionnaire designed by the World Health Organization (WHO) to identify complaints and limitations in children.

This questionnaire consists of two parts. The first part is about body functions. The second part is about the limitations experienced by your son/daughter. Please fill out for every question whether this is a problem to your child or not. If you reply positively, please indicate the burden caused by this complaint/disability. 1 indicates a very little burden, whereas 10 indicates a very high burden.

Does your son/daughter have problems...		Burden due to this problem
...being alert and awake?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with tasks requiring thinking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...falling asleep or staying asleep?	Yes/no	1 2 3 4 5 6 7 8 9 10
...paying attention to someone?	Yes/no	1 2 3 4 5 6 7 8 9 10
...remembering or recalling something?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with clumsiness or coordinating parts of the body?	Yes/no	1 2 3 4 5 6 7 8 9 10
...distinguishing sounds, shapes or smells?	Yes/no	1 2 3 4 5 6 7 8 9 10
...seeing things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...hearing sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...experience pain more than other children of the same age?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the heart?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with allergies or hypersensitivities to any food, plant or animal?	Yes/no	1 2 3 4 5 6 7 8 9 10
...breathing?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with bowel movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...getting the right amount of nutrients?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with urination ?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with the onset menstruation?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with movement of wrists, elbows, shoulders or knees?	Yes/no	1 2 3 4 5 6 7 8 9 10

...with muscles of the body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with stiffness of body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with automatic muscle reflexes?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with balance or body control?	Yes/no	1 2 3 4 5 6 7 8 9 10
...controlling arm and leg movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...tics, tremors or other unusual movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the skin?	Yes/no	1 2 3 4 5 6 7 8 9 10

### **Does your son/daughter have problems...**

### **Burden of this limitation**

...mouthing, touching, or tasting things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing with things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using words, phrases or sentences?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with concepts such as amount, length, the same or different?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing a single task or responding to a single command?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding what others say?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding the meaning of gestures or pictures?	Yes/no	1 2 3 4 5 6 7 8 9 10
...speaking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...making different vocal sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...gestures, pictures or drawings to communicate?	Yes/no	1 2 3 4 5 6 7 8 9 10
...sitting up or getting to stand?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with uncontrolled movements of arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands, fingers and thumb?	Yes/no	1 2 3 4 5 6 7 8 9 10
...walking	Yes/no	1 2 3 4 5 6 7 8 9 10
...relating to others?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in pre-school education?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing alone or with others?	Yes/no	1 2 3 4 5 6 7 8 9 10

## Checklist International Classification of Functioning (3–6 years)

This is a questionnaire designed by the World Health Organization (WHO) to identify complaints and limitations in children.

This questionnaire consists of two parts. The first part is about body functions. The second part is about the limitations experienced by your son/daughter. Please fill out for every question whether this is a problem to your child or not. If you reply positively, please indicate the burden caused by this complaint/disability. 1 indicates a very little burden, whereas 10 indicates a very high burden.

Does your son/daughter have problems...		Burden due to this problem
...being alert and awake?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with tasks requiring thinking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...falling asleep or staying asleep?	Yes/no	1 2 3 4 5 6 7 8 9 10
...paying attention to someone?	Yes/no	1 2 3 4 5 6 7 8 9 10
...remembering or recalling something?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with clumsiness or coordinating parts of the body?	Yes/no	1 2 3 4 5 6 7 8 9 10
...distinguishing sounds, shapes or smells?	Yes/no	1 2 3 4 5 6 7 8 9 10
...seeing things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...hearing sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...experience pain more than other children of the same age?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the heart?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with allergies or hypersensitivities to any food, plant or animal?	Yes/no	1 2 3 4 5 6 7 8 9 10
...breathing?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with bowel movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...getting the right amount of nutrients?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with urination ?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with the onset menstruation?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with movement of wrists, elbows, shoulders or knees?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with muscles of the body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with stiffness of body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with automatic muscle reflexes?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with balance or body control?	Yes/no	1 2 3 4 5 6 7 8 9 10
...controlling arm and leg movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...tics, tremors or other unusual movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the skin?	Yes/no	1 2 3 4 5 6 7 8 9 10

Does your son/daughter have problems...		Burden of this limitation
...mouthing, touching, or tasting things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing with things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using words, phrases or sentences?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with concepts such as amount, length, the same or different?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing a single task or responding to a single command?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding what others say?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding the meaning of gestures or pictures?	Yes/no	1 2 3 4 5 6 7 8 9 10
...speaking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...making different vocal sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...gestures, pictures or drawings to communicate?	Yes/no	1 2 3 4 5 6 7 8 9 10
...sitting up or getting to stand?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with uncontrolled movements of arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands, fingers and thumb?	Yes/no	1 2 3 4 5 6 7 8 9 10
...walking	Yes/no	1 2 3 4 5 6 7 8 9 10
...relating to others?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in pre-school education?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing alone or with others?	Yes/no	1 2 3 4 5 6 7 8 9 10



## Checklist International Classification of Functioning (7-12 years)

This is a questionnaire designed by the World Health Organization (WHO) to identify complaints and limitations in children.

This questionnaire consists of two parts. The first part is about body functions. The second part is about the limitations experienced by your son/daughter. Please fill out for every question whether this is a problem to your child or not. If you reply positively, please indicate the burden caused by this complaint/disability. 1 indicates a very little burden, whereas 10 indicates a very high burden.

Does your son/daughter have problems...		Burden due to this problem
...being alert and awake?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with tasks requiring thinking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...falling asleep or staying asleep?	Yes/no	1 2 3 4 5 6 7 8 9 10
...paying attention to someone?	Yes/no	1 2 3 4 5 6 7 8 9 10
...remembering or recalling something?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with clumsiness or coordinating parts of the body?	Yes/no	1 2 3 4 5 6 7 8 9 10
...distinguishing sounds, shapes or smells?	Yes/no	1 2 3 4 5 6 7 8 9 10
...seeing things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...hearing sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...experience pain more than other children of the same age?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the heart?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with allergies or hypersensitivities to any food, plant or animal?	Yes/no	1 2 3 4 5 6 7 8 9 10
...breathing?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with bowel movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...getting the right amount of nutrients?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with urination ?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with movement of wrists, elbows, shoulders or knees?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with muscles of the body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with stiffness of body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with automatic muscle reflexes?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with balance or body control?	Yes/no	1 2 3 4 5 6 7 8 9 10
...controlling arm and leg movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...tics, tremors or other unusual movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the skin?	Yes/no	1 2 3 4 5 6 7 8 9 10

Does your son/daughter have problems...		Burden of this limitation
...mouthing, touching, or tasting things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing with things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using words, phrases or sentences?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with concepts such as amount, length, the same or different?	Yes/no	1 2 3 4 5 6 7 8 9 10
...learning to read?	Yes/no	1 2 3 4 5 6 7 8 9 10
...learning to write?	Yes/no	1 2 3 4 5 6 7 8 9 10
...learning to calculate?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing a single task or responding to a single command?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing multiple tasks or responding to a command with multiple components?	Yes/no	1 2 3 4 5 6 7 8 9 10
...following the requirements of a daily routine?	Yes/no	1 2 3 4 5 6 7 8 9 10
...being consistent in behaviour from day to day?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding what others say?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding the meaning of gestures or pictures?	Yes/no	1 2 3 4 5 6 7 8 9 10
...speaking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...making different vocal sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...gestures, pictures or drawings to communicate?	Yes/no	1 2 3 4 5 6 7 8 9 10
...sitting up or getting to stand?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with uncontrolled movements of arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands, fingers and thumb?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands?	Yes/no	1 2 3 4 5 6 7 8 9 10
...walking	Yes/no	1 2 3 4 5 6 7 8 9 10
...washing self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using the toilet?	Yes/no	1 2 3 4 5 6 7 8 9 10
...dressing self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...avoiding harm to self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in the activities of the household?	Yes/no	1 2 3 4 5 6 7 8 9 10
...playing alone or with others?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in school education?	Yes/no	1 2 3 4 5 6 7 8 9 10
...engaging in activities in school, neighbourhood or community?	Yes/no	1 2 3 4 5 6 7 8 9 10

## Checklist International Classification of Functioning (over 13 years)

This is a questionnaire designed by the World Health Organization (WHO) to identify complaints and limitations in children.

This questionnaire consists of two parts. The first part is about body functions. The second part is about the limitations experienced by your son/daughter. Please fill out for every question whether this is a problem to your child or not. If you reply positively, please indicate the burden caused by this complaint/disability. 1 indicates a very little burden, whereas 10 indicates a very high burden.

Does your son/daughter have problems...		Burden due to this problem
...being alert and awake?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with tasks requiring thinking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...falling asleep or staying asleep?	Yes/no	1 2 3 4 5 6 7 8 9 10
...paying attention to someone?	Yes/no	1 2 3 4 5 6 7 8 9 10
...remembering or recalling something?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with clumsiness or coordinating parts of the body?	Yes/no	1 2 3 4 5 6 7 8 9 10
...distinguishing sounds, shapes or smells?	Yes/no	1 2 3 4 5 6 7 8 9 10
...seeing things?	Yes/no	1 2 3 4 5 6 7 8 9 10
...hearing sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...experience pain more than other children of the same age?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the heart?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with allergies or hypersensitivities to any food, plant or animal?	Yes/no	1 2 3 4 5 6 7 8 9 10
...breathing?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with bowel movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...getting the right amount of nutrients?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with urination ?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with movement of wrists, elbows, shoulders or knees?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with muscles of the body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with stiffness of body, arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with automatic muscle reflexes?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with balance or body control?	Yes/no	1 2 3 4 5 6 7 8 9 10
...controlling arm and leg movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...tics, tremors or other unusual movements?	Yes/no	1 2 3 4 5 6 7 8 9 10
...the skin?	Yes/no	1 2 3 4 5 6 7 8 9 10

Does your son/daughter have problems...		Burden of this limitation
...reading?	Yes/no	1 2 3 4 5 6 7 8 9 10
...writing?	Yes/no	1 2 3 4 5 6 7 8 9 10
...calculating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...solving problems?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing a single task or responding to a single command?	Yes/no	1 2 3 4 5 6 7 8 9 10
...performing multiple tasks or responding to a command with multiple components?	Yes/no	1 2 3 4 5 6 7 8 9 10
...following the requirements of a daily routine?	Yes/no	1 2 3 4 5 6 7 8 9 10
...being consistent in behaviour from day to day?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding what others say?	Yes/no	1 2 3 4 5 6 7 8 9 10
...understanding the meaning of gestures or pictures?	Yes/no	1 2 3 4 5 6 7 8 9 10
...speaking?	Yes/no	1 2 3 4 5 6 7 8 9 10
...making different vocal sounds?	Yes/no	1 2 3 4 5 6 7 8 9 10
...gestures, pictures or drawings to communicate?	Yes/no	1 2 3 4 5 6 7 8 9 10
...sitting up or getting to stand?	Yes/no	1 2 3 4 5 6 7 8 9 10
...with uncontrolled movements of arms or legs?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands, fingers and thumb?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using hands?	Yes/no	1 2 3 4 5 6 7 8 9 10
...walking	Yes/no	1 2 3 4 5 6 7 8 9 10
...washing self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...using the toilet?	Yes/no	1 2 3 4 5 6 7 8 9 10
...dressing self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...eating?	Yes/no	1 2 3 4 5 6 7 8 9 10
...avoiding harm to self?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in the activities of the household?	Yes/no	1 2 3 4 5 6 7 8 9 10
...in forming and keeping social relationships?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in school education?	Yes/no	1 2 3 4 5 6 7 8 9 10
...participating in programs to prepare for employment?	Yes/no	1 2 3 4 5 6 7 8 9 10
...engaging in activities in school, neighbourhood or community?	Yes/no	1 2 3 4 5 6 7 8 9 10

## Supplementary Document 4

Systematic overview of the outcome measures with good face validity for mitochondrial diseases.

Online at <http://onlinelibrary.wiley.com/store/10.1111/dmcn.12119/asset/supinfo/dmcn12119-sup-0005-FileS5.pdf?v=1&s=deed31905ca20b2566e6eb0bd-3c05d00134bef71>

## References

- Albers CA, Grieve AJ. Test Review: Bayley Scales of Infant and Toddler Development-Third Edition. *Journal of Psychoeducational Assessment*. 2007;2(25):180-90.
- Alhusaini AA, Dean CM, Crosbie J, Shepherd RB, Lewis J. Evaluation of spasticity in children with cerebral palsy using Ashworth and Tardieu Scales compared with laboratory measures. *Journal of child neurology*. 2010 Oct;25(10):1242-7.
- Allison Bender H, Auciello D, Morrison CE, MacAllister WS, Zaroff CM. Comparing the convergent validity and clinical utility of the Behavior Assessment System for Children-Parent Rating Scales and Child Behavior Checklist in children with epilepsy. *Epilepsy & behavior : E&B*. 2008 Jul;13(1):237-42.
- Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW. Underestimation of developmental delay by the new Bayley-III Scale. *Archives of paediatrics & adolescent medicine*. 2010 Apr;164(4):352-6.
- Anderson VB, McKenzie JA, Seton C, Fitzgerald DA, Webster RI, North KN, *et al*. Sniff nasal inspiratory pressure and sleep disordered breathing in childhood neuromuscular disorders. *Neuromuscul Disord*. 2012 Jun;22(6):528-33.
- Angelini C. The role of corticosteroids in muscular dystrophy: a critical appraisal. *Muscle & nerve*. 2007 Oct;36(4):424-35.
- Ashworth B. Preliminary Trial of Carisoprodol in Multiple Sclerosis. *Practitioner*. 1964 Apr;192:540-2.
- Backman E, Henriksson KG. Low-dose prednisolone treatment in Duchenne and Becker muscular dystrophy. *Neuromuscul Disord*. 1995 May;5(3):233-41.
- Baiardini I, Minetti C, Bonifacino S, Porcu A, Klersy C, Petralia P, *et al*. Quality of life in Duchenne muscular dystrophy: the subjective impact on children and parents. *Journal of child neurology*. 2011 Jun;26(6):707-13.
- Baltzopoulos V, Brodie DA. Isokinetic dynamometry. Applications and limitations. *Sports medicine (Auckland, NZ)*. 1989 Aug;8(2):101-16.
- Banugaria SG, Prater SN, Ng YK, Kobori JA, Finkel RS, Ladda RL, *et al*. The impact of antibodies on clinical outcomes in diseases treated with therapeutic protein: lessons learned from infantile Pompe disease. *Genetics in medicine : official journal of the American College of Medical Genetics*. 2011 Aug;13(8):729-36.
- Barry MJ, VanSwearingen JM, Albright AL. Reliability and responsiveness of the Barry-Albright Dystonia Scale. *Developmental medicine and child neurology*. 1999 Jun;41(6):404-11.
- Battini R, Guzzetta A, Sgandurra G, Di Pietro R, Petacchi E, Mercuri E, *et al*. Scale for evaluation of movement disorders in the first three years of life. *Paediatr Neurol*. 2009 Apr;40(4):258-64.
- Battini R, Sgandurra G, Petacchi E, Guzzetta A, Di Pietro R, Giannini MT, *et al*. Movement disorder-childhood rating scale: reliability and validity. *Paediatr Neurol*. 2008 Oct;39(4):259-65.
- Bayley N. Bayley Scales of Infant and Toddler Development-Third Edition: Administration Manual. San Antonio, TX: PsychCorp; 2006.
- Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess*. 1996 Dec;67(3):588-97.
- Beenakker EA, Maurits NM, Fock JM, Brouwer OF, van der Hoeven JH. Functional ability and muscle force in healthy children and ambulant Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol*. 2005;9(6):387-93.
- Beenakker EA, van der Hoeven JH, Fock JM, Maurits NM. Reference values of maximum isometric muscle force obtained in 270 children aged 4-16 years by hand-held dynamometry. *Neuromuscul*

- Disord. 2001 Jul;11(5):441-6.
- Beets MW, Patton MM, Edwards S. The accuracy of pedometer steps and time during walking in children. *Medicine and science in sports and exercise*. 2005 Mar;37(3):513-20.
- Berard C, Payan C, Fermanian J, Girardot F. [A motor function measurement scale for neuromuscular diseases - description and validation study]. *Rev Neurol (Paris)*. 2006 Apr;162(4):485-93.
- Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord*. 2005 Jul;15(7):463-70.
- Biering-Sorensen F, Nielsen JB, Klinge K. Spasticity-assessment: a review. *Spinal Cord*. 2006 Dec;44(12):708-22.
- Bjornson KF. Physical activity monitoring in children and youths. *Paediatr Phys Ther*. 2005 Spring;17(1):37-45.
- Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther*. 1987 Feb;67(2):206-7.
- Boiteau M, Malouin F, Richards CL. Use of a hand-held dynamometer and a Kin-Com dynamometer for evaluating spastic hypertonia in children: a reliability study. *Phys Ther*. 1995 Sep;75(9):796-802.
- Bourke-Taylor H. Melbourne Assessment of Unilateral Upper Limb Function: construct validity and correlation with the Pediatric Evaluation of Disability Inventory. *Developmental medicine and child neurology*. 2003 Feb;45(2):92-6.
- Brunton LK, Bartlett DJ. Validity and reliability of two abbreviated versions of the Gross Motor Function Measure. *Phys Ther*. 2011 Apr;91(4):577-88.
- Buchalter MB, Weiss JL, Rogers WJ, Zerhouni EA, Weisfeldt ML, Beyar R, *et al*. Noninvasive quantification of left ventricular rotational deformation in normal humans using magnetic resonance imaging myocardial tagging. *Circulation*. 1990 Apr;81(4):1236-44.
- Burk K, Malzig U, Wolf S, Heck S, Dimitriadis K, Schmitz-Hubsch T, *et al*. Comparison of three clinical rating scales in Friedreich ataxia (FRDA). *Movement disorders : official journal of the Movement Disorder Society*. 2009 Sep 15;24(12):1779-84.
- Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology*. 1985 Jan;35(1):73-7.
- Campbell SK, Kolobe TH, Wright BD, Linacre JM. Validity of the Test of Infant Motor Performance for prediction of 6-, 9- and 12-month scores on the Alberta Infant Motor Scale. *Developmental medicine and child neurology*. 2002 Apr;44(4):263-72.
- Cano SJ, Hobart JC, Hart PE, Korlipara LV, Schapira AH, Cooper JM. International Cooperative Ataxia Rating Scale (ICARS): appropriate for studies of Friedreich's ataxia? *Movement disorders : official journal of the Movement Disorder Society*. 2005 Dec;20(12):1585-91.
- Capio CM, Sit CH, Abernethy B. Physical activity measurement using MTI (actigraph) among children with cerebral palsy. *Arch Phys Med Rehabil*. 2010 Aug;91(8):1283-90.
- Capio CM, Sit CH, Abernethy B, Rotor ER. Physical activity measurement instruments for children with cerebral palsy: a systematic review. *Developmental medicine and child neurology*. 2010 Oct;52(10):908-16.
- Carter T, Callaghan P, Khalil E, Morris I. The effectiveness of a preferred intensity exercise programme on the mental health outcomes of young people with depression: a sequential mixed methods evaluation. *BMC public health*. 2012;12:187.
- Chen TH, Chang JG, Yang YH, Mai HH, Liang WC, Wu YC, *et al*. Randomised, double-blind, placebo-controlled trial of hydroxyurea in spinal muscular atrophy. *Neurology*. 2010 Dec 14;75(24):2190-7.
- Clanchy KM, Tweedy SM, Boyd R. Measurement of habitual physical activity performance in adoles-

cents with cerebral palsy: a systematic review. *Developmental medicine and child neurology*. 2011 Jun;53(6):499-505.

Clanchy KM, Tweedy SM, Boyd RN, Trost SG. Validity of accelerometry in ambulatory children and adolescents with cerebral palsy. *European journal of applied physiology*. 2011 Dec;111(12):2951-9.

Connolly BH, Dalton L, Smith JB, Lamberth NG, McCay B, Murphy W. Concurrent validity of the Bayley Scales of Infant Development II (BSID-II) Motor Scale and the Peabody Developmental Motor Scale II (PDMS-2) in 12-month-old infants. *Paediatr Phys Ther*. 2006 Fall;18(3):190-6.

Council MR. Aids to the investigation of peripheral nerve injuries. London: HMSO; 1976.

Custers JW, Wassenberg-Severijnen JE, Van der Net J, Vermeer A, Hart HT, Helders PJ. Dutch adaptation and content validity of the 'Pediatric Evaluation Of Disability Inventory (PEDI)'. *Disabil Rehabil*. 2002 Mar 20;24(5):250-8.

Damiano DL, Quinlivan JM, Owen BF, Payne P, Nelson KC, Abel MF. What does the Ashworth scale really measure and are instrumented measures more valid and precise? *Developmental medicine and child neurology*. 2002 Feb;44(2):112-8.

Darrah J, Piper M, Watt MJ. Assessment of gross motor skills of at-risk infants: predictive validity of the Alberta Infant Motor Scale. *Developmental medicine and child neurology*. 1998 Jul;40(7):485-91.

Davis E, Shelly A, Waters E, Davern M. Measuring the quality of life of children with cerebral palsy: comparing the conceptual differences and psychometric properties of three instruments. *Developmental medicine and child neurology*. 2010 Feb;52(2):174-80.

Davis MF. Measuring impairment and functional limitations in children with cerebral palsy. *Disabil Rehabil*. 2011;33(25-26):2416-24.

Davis SE, Hyman LS, Limbers CA, Andersen CM, Greene MC, Varni JW, *et al*. The PedsQL in paediatric patients with Duchenne muscular dystrophy: feasibility, reliability, and validity of the Paediatric Quality of Life Inventory Neuromuscular Module and Generic Core Scales. *J Clin Neuromuscul Dis*. 2010 Mar;11(3):97-109.

de Groot JF, Takken T, van Brussel M, Gooskens R, Schoenmakers M, Versteeg C, *et al*. Randomised controlled study of home-based treadmill training for ambulatory children with spina bifida. *Neurorehabil Neural Repair*. 2011 Sep;25(7):597-606.

Debusse D, Brace H. Outcome measures of activity for children with cerebral palsy: a systematic review. *Paediatr Phys Ther*. 2011 Fall;23(3):221-31.

Delatycki MB. Evaluating the progression of Friedreich ataxia and its treatment. *J Neurol*. 2009 Mar;256 Suppl 1:36-41.

deMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. QUEST: Quality of Upper Extremity Skills Test. Hamilton, ON: McMaster University, Neurodevelopmental Clinical Research Unit; 1992.

DeMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The reliability and validity of Quality of Upper Extremity Skills Test. *Phys Occup Ther Paediatr*. 1993;13(2):1-18.

Donders J. Validity of the Kaufman Assessment Battery for Children when employed with children with traumatic brain injury. *Journal of clinical psychology*. 1992 Mar;48(2):225-30.

Duan X, Zhang S, Xiao N. Reliability and validity of the PedsQL Generic Core Scales 4.0 for Chinese children with epilepsy. *Epilepsy & behavior* : E&B. 2012 Apr;23(4):431-6.

Edwards S, Fletcher P, Gurman M, Hughes A, Letts C. Reynell Developmental Language Scales III (RDLS III). The University of Reading Edition ed: NFER-Nelson; 1997.

Eugster-Buesch F, de Bruin ED, Boltshauser E, Steinlin M, Kuenzle C, Muller E, *et al*. Forced-use therapy for children with cerebral palsy in the community setting: A single-blinded randomised



- controlled pilot trial. *Journal of paediatric rehabilitation medicine*. 2012 Jan 1;5(2):65-74.
- Eyberg S. *Eyberg Child Behavior Inventory*. Lutz, Florida: Psychological Assessment Resources; 1999.
- Eyberg S, Ross A. Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child Psychology*. 1978;113-6.
- Fauroux B, Aubertin G, Cohen E, Clement A, Lofaso F. Sniff nasal inspiratory pressure in children with muscular, chest wall or lung disease. *Eur Respir J*. 2009 Jan;33(1):113-7.
- Febrer A, Vigo M, Fagoaga J, Medina-Cantillo J, Rodriguez N, Tizzano E. [Hammersmith functional rating scale for children with spinal muscular atrophy. Validation of the Spanish version]. *Revista de neurologia*. 2011 Dec 1;53(11):657-63.
- Fee RJ, Hinton VJ. Resilience in children diagnosed with a chronic neuromuscular disorder. *J Dev Behav Paediatr*. 2011 Nov;32(9):644-50.
- Ferrario SR, Zotti AM, Baroni A, Cavagnino A, Fornara R. Emotional reactions and practical problems of the caregivers of hemodialysed patients. *Journal of nephrology*. 2002 Jan-Feb;15(1):54-60.
- Figueras Masip A, Amador-Campos JA, Gomez-Benito J, del Barrio Gandara V. Psychometric properties of the Children's Depression Inventory in community and clinical sample. *Span J Psychol*. 2010 Nov;13(2):990-9.
- Fleischman A, Makimura H, Stanley TL, McCarthy MA, Kron M, Sun N, *et al*. Skeletal muscle phosphocreatine recovery after submaximal exercise in children and young and middle-aged adults. *J Clin Endocrinol Metab*. 2010 Sep;95(9):E69-74.
- Florence JM, Pandya S, King WM, Robison JD, Signore LC, Wentzell M, *et al*. Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. *Phys Ther*. 1984 Jan;64(1):41-5.
- Folio M.R., R.R. F. *Peabody Developmental Motor Scales Second Edition*. Austin Texas: Pro-Ed; 2000.
- Gauld LM, Boynton A, Betts GA, Johnston H. Spirometry is affected by intelligence and behavior in Duchenne muscular dystrophy. *Paediatr Pulmonol*. 2005 Nov;40(5):408-13.
- Gimeno H, Tustin K, Selway R, Lin JP. Beyond the Burke-Fahn-Marsden Dystonia Rating Scale: Deep brain stimulation in childhood secondary dystonia. *Eur J Paediatr Neurol*. 2012 Jan 16.
- Gordijn SM, Cremers EM, Kaspers GJ, Gemke RJ. Fatigue in children: reliability and validity of the Dutch PedsQL(TM) Multidimensional Fatigue Scale. *Qual Life Res*. 2011 Jan 19;20(Sep):1103-8.
- Gracies JM, Burke K, Clegg NJ, Browne R, Rushing C, Fehlings D, *et al*. Reliability of the Tardieu Scale for assessing spasticity in children with cerebral palsy. *Arch Phys Med Rehabil*. 2010 Mar;91(3):421-8.
- Gross D, Fogg L, Young M, Ridge A, Cowell J, Sivan A, *et al*. Reliability and validity of the Eyberg Child Behavior Inventory with African-American and Latino parents of young children. *Research in nursing & health*. 2007 Apr;30(2):213-23.
- Hebert LJ, Maltais DB, Lepage C, Saulnier J, Crete M, Perron M. Isometric muscle strength in youth assessed by hand-held dynamometry: a feasibility, reliability, and validity study. *Paediatr Phys Ther*. 2011 Fall;23(3):289-99.
- Hebestreit H. Exercise testing in children -- what works, what doesn't, and where to go? *Paediatric respiratory reviews*. 2004;5 Suppl A:S11-4.
- Heritier F, Rahm F, Pasche P, Fitting JW. Sniff nasal inspiratory pressure. A noninvasive assessment of inspiratory muscle strength. *American journal of respiratory and critical care medicine*. 1994 Dec;150(6 Pt 1):1678-83.

Hicks JE, Drinkard B, Summers RM, Rider LG. Decreased aerobic capacity in children with juvenile dermatomyositis. *Arthritis and rheumatism*. 2002 Apr 15;47(2):118-23.

Hiller LB, Wade CK. Upper extremity functional assessment scales in children with Duchenne muscular dystrophy: a comparison. *Arch Phys Med Rehabil*. 1992 Jun;73(6):527-34.

Hoare BJ, Wasiak J, Imms C, Carey L. Constraint-induced movement therapy in the treatment of the upper limb in children with hemiplegic cerebral palsy. *Cochrane Database Syst Rev*. 2007(2):CD004149.

Hollingsworth KG, Gorman GS, Trenell MI, McFarland R, Taylor RW, Turnbull DM, *et al*. Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load. *Neuromuscul Disord*. 2012 Jul;22(7):592-6.

Houltram J, Noble I, Boyd RN, Corry I, Flett P, Graham HK. Botulinum toxin type A in the management of equinus in children with cerebral palsy: an evidence-based economic evaluation. *Eur J Neurol*. 2001 Nov;8 Suppl 5:194-202.

Huer MB, Miller L. *Test of Early Communication and Emerging Language*. Austin, Texas.: Pro-ed; 2011.

Hullmann SE, Ryan JL, Ramsey RR, Chaney JM, Mullins LL. Measures of general paediatric quality of life: Child Health Questionnaire (CHQ), DISABKIDS Chronic Generic Measure (DCGM), KINDL-R, Paediatric Quality of Life Inventory (PedsQL) 4.0 Generic Core Scales, and Quality of My Life Questionnaire (QoML). *Arthritis care & research*. 2011 Nov;63 Suppl 11:S420-30.

Iannaccone ST, Hynan LS. Reliability of 4 outcome measures in paediatric spinal muscular atrophy. *Arch Neurol*. 2003 Aug;60(8):1130-6.

Jansen M, de Groot IJ, van Alfen N, Geurts A. Physical training in boys with Duchenne Muscular Dystrophy: the protocol of the No Use is Disuse study. *BMC paediatrics*. 2010;10:55.

Jeannet PY, Aminian K, Bloetzer C, Najafi B, Paraschiv-Ionescu A. Continuous monitoring and quantification of multiple parameters of daily physical activity in ambulatory Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol*. 2011 Jan;15(1):40-7.

Jebsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. *Arch Phys Med Rehabil*. 1969 Jun;50(6):311-9.

Jeng SF, Yau KI, Chen LC, Hsiao SF. Alberta infant motor scale: reliability and validity when used on preterm infants in Taiwan. *Phys Ther*. 2000 Feb;80(2):168-78.

Jeppesen TD, Schwartz M, Frederiksen AL, Wibrand F, Olsen DB, Vissing J. Muscle phenotype and mutation load in 51 persons with the 3243A>G mitochondrial DNA mutation. *Arch Neurol*. 2006 Dec;63(12):1701-6.

Johnson LM, Randall MJ, Reddihough DS, Oke LE, Byrt TA, Bach TM. Development of a clinical assessment of quality of movement for unilateral upper-limb function. *Developmental medicine and child neurology*. 1994 Nov;36(11):965-73.

Johnson TS, Engstrom JL, Gelhar DK. Intra- and interexaminer reliability of anthropometric measurements of term infants. *J Paediatr Gastroenterol Nutr*. 1997;24(5):497-505.

Katz-Leurer M, Rotem H, Keren O, Meyer S. The relationship between step variability, muscle strength and functional walking performance in children with post-traumatic brain injury. *Gait & posture*. 2009 Jan;29(1):154-7.

Kaufmann P, McDermott MP, Darras BT, Finkel R, Kang P, Oskoui M, *et al*. Observational study of spinal muscular atrophy type 2 and 3: functional outcomes over 1 year. *Arch Neurol*. 2011 Jun;68(6):779-86.

- Kaufmann P, Shungu DC, Sano MC, Jhung S, Engelstad K, Mitsis E, *et al.* Cerebral lactic acidosis correlates with neurological impairment in MELAS. *Neurology*. 2004 Apr 27;62(8):1297-302.
- Kearney M, Orrell RW, Fahey M, Pandolfo M. Antioxidants and other pharmacological treatments for Friedreich ataxia. *Cochrane Database Syst Rev*. 2012;4:CD007791.
- Ketelslegers IA, Catsman-Berrevoets CE, Boon M, Eikelenboom MJ, Stroink H, Neuteboom RF, *et al.* Fatigue and depression in children with multiple sclerosis and monophasic variants. *Eur J Paediatr Neurol*. 2010 Jul;14(4):320-5.
- King GA, Law M, King S, Hurley P, Hanna S, Kertoy M, *et al.* Measuring children's participation in recreation and leisure activities: construct validation of the CAPE and PAC. *Child: care, health and development*. 2007 Jan;33(1):28-39.
- King GL, M.; King S.; *et al.* Children's assessment of participation and enjoyment & preferences for activities of children: San Antonio: Hartcourt Assessment; 2004.
- Klepper SE. Measures of paediatric function: Child Health Assessment Questionnaire (C-HAQ), Juvenile Arthritis Functional Assessment Scale (JAFAS), Paediatric Outcomes Data Collection Instrument (PODCI), and Activities Scale for Kids (ASK). *Arthritis care & research*. 2011 Nov;63 Suppl 11:S371-82.
- Klinge L, Straub V, Neudorf U, Voit T. Enzyme replacement therapy in classical infantile pompe disease: results of a ten-month follow-up study. *Neuropediatrics*. 2005 Feb;36(1):6-11.
- Klingels K, De Cock P, Desloovere K, Huenaerts C, Molenaers G, Van Nuland I, *et al.* Comparison of the Melbourne Assessment of Unilateral Upper Limb Function and the Quality of Upper Extremity Skills Test in hemiplegic CP. *Developmental medicine and child neurology*. 2008 Dec;50(12):904-9.
- Kornblum C, Schroder R, Muller K, Vorgerd M, Eggers J, Bogdanow M, *et al.* Creatine has no beneficial effect on skeletal muscle energy metabolism in patients with single mitochondrial DNA deletions: a placebo-controlled, double-blind 31P-MRS cross-over study. *Eur J Neurol*. 2005 Apr;12(4):300-9.
- Koskentausta T, Iivanainen M, Almqvist F. CBCL in the assessment of psychopathology in Finnish children with intellectual disability. *Res Dev Disabil*. 2004 Jul-Aug;25(4):341-54.
- Krystkowiak P, du Montcel ST, Vercueil L, Houeto JL, Lagrange C, Cornu P, *et al.* Reliability of the Burke-Fahn-Marsden scale in a multicentre trial for dystonia. *Movement disorders : official journal of the Movement Disorder Society*. 2007 Apr 15;22(5):685-9.
- Kwon DG, Chung CY, Lee KM, Lee DJ, Lee SC, Choi IH, *et al.* Transcultural adaptation and validation of the Korean version of the Paediatric Outcomes Data Collection Instrument (PODCI) in children and adolescents. *Journal of paediatric orthopedics*. 2011 Jan-Feb;31(1):102-6.
- Larson-Meyer DE, Newcomer BR, Hunter GR, Hetherington HP, Weinsier RL. 31P MRS measurement of mitochondrial function in skeletal muscle: reliability, force-level sensitivity and relation to whole body maximal oxygen uptake. *NMR in biomedicine*. 2000 Jan;13(1):14-27.
- Laufersweiler-Plass C, Rudnik-Schoneborn S, Zerres K, Backes M, Lehmkuhl G, von Gontard A. Behavioural problems in children and adolescents with spinal muscular atrophy and their siblings. *Developmental medicine and child neurology*. 2003 Jan;45(1):44-9.
- Launey KB, Carroll J, Van Horn KR. Concurrent validity of the WISC-IV in eligibility decisions for students with educable mental disabilities. *Psychological reports*. 2007 Jun;100(3 Pt 2):1165-70.
- Law M, Cadman D, Rosenbaum P, Walter S, Russell D, DeMatteo C. Neurodevelopmental therapy and upper-extremity inhibitive casting for children with cerebral palsy. *Developmental medicine and child neurology*. 1991 May;33(5):379-87.

Law M, King G, King S, Kertoy M, Hurley P, Rosenbaum P, *et al.* Patterns of participation in recreational and leisure activities among children with complex physical disabilities. *Developmental medicine and child neurology*. 2006 May;48(5):337-42.

Law MC, Darrah J, Pollock N, Wilson B, Russell DJ, Walter SD, *et al.* Focus on function: a cluster, randomised controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental medicine and child neurology*. 2011 Jul;53(7):621-9.

Lee HFT, C.R., Chi, C.S., Lee, H.J., Chen, C.C. Leigh syndrome: Clinical and Neuroimaging Follow-up. *Paediatric Neurology*. 2009;40(2):88-93.

Lee SK, Kim J, Kim HD, Lee JS, Lee YM. Initial experiences with proton MR spectroscopy in treatment monitoring of mitochondrial encephalopathy. *Yonsei Med J*. 2010 Sep;51(5):672-5.

Lerman JA, Sullivan E, Haynes RJ. The Paediatric Outcomes Data Collection Instrument (PODCI) and functional assessment in patients with adolescent or juvenile idiopathic scoliosis and congenital scoliosis or kyphosis. *Spine*. 2002 Sep 15;27(18):2052-7; discussion 7-8.

Lesser DJ, Fleming MM, Maher CA, Kim SB, Woo MS, Keens TG. Does the 6-min walk test correlate with the exercise stress test in children? *Paediatr Pulmonol*. 2010 Feb;45(2):135-40.

Li AM, Yin J, Yu CC, Tsang T, So HK, Wong E, *et al.* The six-minute walk test in healthy children: reliability and validity. *Eur Respir J*. 2005 Jun;25(6):1057-60.

Lynch DR, Perlman SL, Meier T. A phase 3, double-blind, placebo-controlled trial of idebenone in friedreich ataxia. *Arch Neurol*. 2010 Aug;67(8):941-7.

Maher CA, Williams MT, Olds T, Lane AE. Physical and sedentary activity in adolescents with cerebral palsy. *Developmental medicine and child neurology*. 2007 Jun;49(6):450-7.

Main M, Kairon H, Mercuri E, Muntoni F. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. *Eur J Paediatr Neurol*. 2003;7(4):155-9.

Majnemer A, Shevell M, Law M, Poulin C, Rosenbaum P. Indicators of distress in families of children with cerebral palsy. *Disabil Rehabil*. 2012 Jan 12.

Malda M, van de Vijver FJ, Srinivasan K, Transler C, Sukumar P. Traveling with cognitive tests: testing the validity of a KABC-II adaptation in India. *Assessment*. 2010 Mar;17(1):107-15.

Manzur AY, Kuntzer T, Pike M, Swan A. Glucocorticoid corticosteroids for Duchenne muscular dystrophy. *Cochrane Database Syst Rev*. 2008(1):CD003725.

Marcus KA, Barends M, Morava-Kozicz E, Feuth T, de Korte CL, Kapusta L. Early detection of myocardial dysfunction in children with mitochondrial disease: An ultrasound and two-dimensional strain echocardiography study. *Mitochondrion*. 2011 Dec 10.

Marcus KA, de Korte CL, Feuth T, Thijssen JM, van Oort AM, Tanke RB, *et al.* Persistent reduction in left ventricular strain using two-dimensional speckle-tracking echocardiography after balloon valvuloplasty in children with congenital valvular aortic stenosis. *J Am Soc Echocardiogr*. 2012 May;25(5):473-85.

Marelli C, Figoni J, Charles P, Anheim M, Tchikviladze M, Vincitorio CM, *et al.* Annual change in Friedreich's ataxia evaluated by the Scale for the Assessment and Rating of Ataxia (SARA) is independent of disease severity. *Movement disorders : official journal of the Movement Disorder Society*. 2012 Jan;27(1):135-8.

Mayes TL, Bernstein IH, Haley CL, Kennard BD, Emslie GJ. Psychometric properties of the Children's Depression Rating Scale-Revised in adolescents. *Journal of child and adolescent psychopharmacology*. 2010 Dec;20(6):513-6.

- Mayhew JE, Florence JM, Mayhew TP, Henricson EK, Leshner RT, McCarter RJ, *et al.* Reliable surrogate outcome measures in multicentre clinical trials of Duchenne muscular dystrophy. *Muscle & nerve*. 2007 Jan;35(1):36-42.
- McCarthy ML, Silberstein CE, Atkins EA, Harryman SE, Sponseller PD, Hadley-Miller NA. Comparing reliability and validity of paediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Developmental medicine and child neurology*. 2002 Jul;44(7):468-76.
- McCullough N, Parkes J. Use of the child health questionnaire in children with cerebral palsy: a systematic review and evaluation of the psychometric properties. *Journal of paediatric psychology*. 2008 Jan Feb;33(1):80-90.
- McDonald CM, Abresch RT, Carter GT, Fowler WM, Jr., Johnson ER, Kilmer DD, *et al.* Profiles of neuromuscular diseases. Duchenne muscular dystrophy. *Am J Phys Med Rehabil*. 1995 Sep-Oct;74(5 Suppl):S70-92.
- McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, *et al.* The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle & nerve*. 2010 Apr;41(4):500-10.
- McDonald CM, Widman L, Abresch RT, Walsh SA, Walsh DD. Utility of a step activity monitor for the measurement of daily ambulatory activity in children. *Arch Phys Med Rehabil*. 2005 Apr;86(4):793-801.
- McDonald CM, Widman LM, Walsh DD, Walsh SA, Abresch RT. Use of step activity monitoring for continuous physical activity assessment in boys with Duchenne muscular dystrophy. *Arch Phys Med Rehabil*. 2005 Apr;86(4):802-8.
- Mercuri E, Messina S, Battini R, Berardinelli A, Boffi P, Bono R, *et al.* Reliability of the Hammersmith functional motor scale for spinal muscular atrophy in a multicentric study. *Neuromuscul Disord*. 2006 Feb;16(2):93-8.
- Merlini L, Bertini E, Minetti C, Mongini T, Morandi L, Angelini C, *et al.* Motor function-muscle strength relationship in spinal muscular atrophy. *Muscle & nerve*. 2004 Apr;29(4):548-52.
- Moller HE, Wiedermann D, Kurlemann G, Hilbich T, Schuierer G. Application of NMR spectroscopy to monitoring MELAS treatment: a case report. *Muscle & nerve*. 2002 Apr;25(4):593-600.
- Monbaliu E, Ortibus E, Roelens F, Desloovere K, Deklerck J, Prinzie P, *et al.* Rating scales for dystonia in cerebral palsy: reliability and validity. *Developmental medicine and child neurology*. 2010 Jun;52(6):570-5.
- Montes J, Gordon AM, Pandya S, De Vivo DC, Kaufmann P. Clinical outcome measures in spinal muscular atrophy. *Journal of child neurology*. 2009 Aug;24(8):968-78.
- Morava E, Gardeitchik T, Kozicz T, de Boer L, Koene S, de Vries MC, *et al.* Depressive behaviour in children diagnosed with a mitochondrial disorder. *Mitochondrion*. 2010 Aug;10(5):528-33.
- Msall ME. Measuring functional skills in preschool children at risk for neurodevelopmental disabilities. *Mental retardation and developmental disabilities research reviews*. 2005;11(3):263-73.
- Msall ME, DiGaudio K, Rogers BT, LaForest S, Catanzaro NL, Campbell J, *et al.* The Functiona Independence Measure for Children (WeeFIM). Conceptual basis and pilot use in children with developmental disabilities. *Clin Paediatr (Phila)*. 1994 Jul;33(7):421-30.
- Muenzer J, Gucsavas-Calikoglu M, McCandless SE, Schuetz TJ, Kimura A. A phase I/II clinical trial of enzyme replacement therapy in mucopolysaccharidosis II (Hunter syndrome). *Molecular genetics and metabolism*. 2007 Mar;90(3):329-37.
- N B. Bayley scales of infant and toddler development. Harcourt Assessment. 2006;3rd ed.
- Nelson L, Owens H, Hyman LS, Iannaccone ST. The gross motor function measure is a valid and sensitive outcome measure for spinal muscular atrophy. *Neuromuscul Disord*. 2006 Jun;16(6):374-80.

Nicot F, Hart N, Forin V, Boule M, Clement A, Polkey MI, *et al.* Respiratory muscle testing: a valuable tool for children with neuromuscular disorders. *American journal of respiratory and critical care medicine*. 2006 Jul 1;174(1):67-74.

Niewczyk PM, Granger CV. Measuring function in young children with impairments. *Paediatr Phys Ther*. 2010 Spring;22(1):42-51.

Nordmark E, Jarnlo GB, Hagglund G. Comparison of the Gross Motor Function Measure and Pediatric Evaluation of Disability Inventory in assessing motor function in children undergoing selective dorsal rhizotomy. *Developmental medicine and child neurology*. 2000 Apr;42(4):245-52.

Olesch CA, Greaves S, Imms C, Reid SM, Graham HK. Repeat botulinum toxin-A injections in the upper limb of children with hemiplegia: a randomised controlled trial. *Developmental medicine and child neurology*. 2010 Jan;52(1):79-86.

Osman A, Barrios FX, Gutierrez PM, Williams JE, Bailey J. Psychometric properties of the Beck Depression Inventory-II in nonclinical adolescent samples. *Journal of clinical psychology*. 2008 Jan;64(1):83-102.

Ottenbacher KJ, Msall ME, Lyon N, Duffy LC, Ziviani J, Granger CV, *et al.* The WeeFIM instrument: its utility in detecting change in children with developmental disabilities. *Arch Phys Med Rehabil*. 2000 Oct;81(10):1317-26.

Ottenbacher KJ, Taylor ET, Msall ME, Braun S, Lane SJ, Granger CV, *et al.* The stability and equivalence reliability of the functional independence measure for children (WeeFIM). *Developmental medicine and child neurology*. 1996 Oct;38(10):907-16.

Pandyan AD, Price CI, Rodgers H, Barnes MP, Johnson GR. Biomechanical examination of a commonly used measure of spasticity. *Clinical biomechanics (Bristol, Avon)*. 2001 Dec;16(10):859-65.

Parreira SL, Resende MB, Della Corte Peduto M, Marie SK, Carvalho MS, Reed UC. Quantification of muscle strength and motor ability in patients with Duchenne muscular dystrophy on steroid therapy. *Arq Neuropsiquiatr*. 2007 Jun;65(2A):245-50.

Parreira SL, Resende MB, Zanoteli E, Carvalho MS, Marie SK, Reed UC. Comparison of motor strength and function in patients with Duchenne muscular dystrophy with or without steroid therapy. *Arq Neuropsiquiatr*. 2010 Oct;68(5):683-8.

Paulsen EK, Friedman LS, Myers LM, Lynch DR. Health-related quality of life in children with Friedreich ataxia. *Paediatr Neurol*. 2010 May;42(5):335-7.

Pavakis SG, Kingsley PB, Kaplan GP, Stacpoole PW, O'Shea M, Lustbader D. Magnetic resonance spectroscopy: use in monitoring MELAS treatment. *Arch Neurol*. 1998 Jun;55(6):849-52.

Pavone L, Burton J, Gaebler-Spira D. Dystonia in Childhood: Clinical and Objective Measures and Functional Implications. *Journal of child neurology*. 2012 Jun 29.

Phoenix C, Schaefer AM, Elson JL, Morava E, Bugiani M, Uziel G, *et al.* A scale to monitor progression and treatment of mitochondrial disease in children. *Neuromuscul Disord*. 2006 Dec;16(12):814-20.

Pierce SR, Lauer RT, Shewokis PA, Rubertone JA, Orlin MN. Test-retest reliability of isokinetic dynamometry for the assessment of spasticity of the knee flexors and knee extensors in children with cerebral palsy. *Arch Phys Med Rehabil*. 2006 May;87(5):697-702.

Piper MC, Pinnell LE, Darrah J, Byrne PJ, Watt MJ. Early developmental screening: sensitivity and specificity of chronological and adjusted scores. *J Dev Behav Paediatr*. 1992 Apr;13(2):95-101.

Poznanski EO, Mokros HB. *Children's Depression Rating Scale, Revised (CDRS-R)*. Los Angeles, CA WPS; 1995.

Raat H, Bonsel GJ, Essink-Bot ML, Landgraf JM, Gemke RJ. Reliability and validity of comprehensive health status measures in children: The Child Health Questionnaire in relation to the Health

- Utilities Index. *Journal of clinical epidemiology*. 2002 Jan;55(1):67-76.
- Raat H, Landgraf JM, Oostenbrink R, Moll HA, Essink-Bot ML. Reliability and validity of the Infant and Toddler Quality of Life Questionnaire (ITQOL) in a general population and respiratory disease sample. *Qual Life Res*. 2007 Apr;16(3):445-60.
- Rackley C, Allen DN, Fuhrman LJ, Mayfield J. Generalizability of WISC-IV index and subtest score profiles in children with traumatic brain injury. *Child neuropsychology : a journal on normal and abnormal development in childhood and adolescence*. 2011 Nov 22.
- Rafferty GF, Leech S, Knight L, Moxham J, Greenough A. Sniff nasal inspiratory pressure in children. *Paediatr Pulmonol*. 2000 Jun;29(6):468-75.
- Raggi A, Leonardi M. Assessing activity limitations in patients with neuromuscular diseases: is the ACTVILIM questionnaire linked to ICF and ICF-CY? *Int J Rehabil Res*. 2009 Jun;32(2):148-53.
- Rahman T, Sample W, Seliktar R, Scavina MT, Clark AL, Moran K, *et al*. Design and testing of a functional arm orthosis in patients with neuromuscular diseases. *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*. 2007 Jun;15(2):244-51.
- Randall M, Carlin JB, Chondros P, Reddihough D. Reliability of the Melbourne assessment of unilateral upper limb function. *Developmental medicine and child neurology*. 2001 Nov;43(11):761-7.
- Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Duer W, *et al*. KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Rev Pharmacoecon Outcomes Res*. 2005 Jun;5(3):353-64.
- Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Power M, *et al*. The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2008 Jul-Aug;11(4):645-58.
- Reynolds C.R., R.W. K. *Behavior Assessment System for Children: Second Edition*. Circle Pines, Minnesota: AGS Publishing 2004.
- Robertson RJ, Goss FL, Aaron DJ, Tessmer KA, Gairola A, Ghigiarelli JJ, *et al*. Observation of perceived exertion in children using the OMNI pictorial scale. *Medicine and science in sports and exercise*. 2006 Jan;38(1):158-66.
- Robertson RJ, Goss FL, Boer NF, Peoples JA, Foreman AJ, Dabayeb IM, *et al*. Children's OMNI scale of perceived exertion: mixed gender and race validation. *Medicine and science in sports and exercise*. 2000 Feb;32(2):452-8.
- Roemmich JN, Barkley JE, Epstein LH, Lobarinas CL, White TM, Foster JH. Validity of PCERT and OMNI walk/run ratings of perceived exertion. *Medicine and science in sports and exercise*. 2006 May;38(5):1014-9.
- Rossi Ferrario S, Baiardi P, Zotti AM. [Assessment of problems associated with caregiving: the family strain questionnaire]. *Giornale italiano di medicina del lavoro ed ergonomia*. 2001 Jan-Mar;23(1):25-9.
- Rossi Ferrario S, Zotti AM, Massara G, Nuvolone G. A comparative assessment of psychological and psychosocial characteristics of cancer patients and their caregivers. *Psychooncology*. 2003 Jan-Feb;12(1):1-7.
- Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L, *et al*. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. *Clinical and experimental rheumatology*. 2001 Jul-Aug;19(4 Suppl 23):S1-9.

Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S. The gross motor function measure: a means to evaluate the effects of physical therapy. *Developmental medicine and child neurology*. 1989 Jun;31(3):341-52.

Russell DJ RP, Avery LM, Lane M. *Gross Motor Function Measure (GMGM-66 and GMFM-88) User's Manual*. London, United Kingdom: Mac Keith Press. 2002.

Ryan JJ, Glass LA, Bartels JM. Stability of the WISC-IV in a sample of elementary and middle school children. *Applied neuropsychology*. 2010 Jan;17(1):68-72.

Sakzewski L, Carlon S, Shields N, Ziviani J, Ware RS, Boyd RN. Impact of intensive upper limb rehabilitation on quality of life: a randomised trial in children with unilateral cerebral palsy. *Developmental medicine and child neurology*. 2012 May;54(5):415-23.

Saylor CF, Finch AJ, Spirito A, Bennett B. Children's Depression Inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology*. 1984;52(6):955-67.

Scholtes VA, Becher JG, Beelen A, Lankhorst GJ. Clinical assessment of spasticity in children with cerebral palsy: a critical review of available instruments. *Developmental medicine and child neurology*. 2006 Jan;48(1):64-73.

Semel E, Wiig EH. *Clinical Evaluation of Language Fundamentals - Fourth Edition* Harlow, Essex: PsychCorp; 2006.

Seo Y, Ishizu T, Enomoto Y, Sugimori H, Yamamoto M, Machino T, *et al*. Validation of 3-dimensional speckle tracking imaging to quantify regional myocardial deformation. *Circulation Cardiovascular imaging*. 2009 Nov;2(6):451-9.

Sherman EM, Brooks BL, Fay-McClymont TB, Macallister WS. Detecting epilepsy-related cognitive problems in clinically referred children with epilepsy: Is the WISC-IV a useful tool? *Epilepsia*. 2012 Jun;53(6):1060-6.

Sival DA, Pouwels ME, Van Brederode A, Maurits NM, Verschuuren-Bemelmans CC, Brunt ER, *et al*. In children with Friedreich ataxia, muscle and ataxia parameters are associated. *Developmental medicine and child neurology*. 2011 Jun;53(6):529-34.

Sorensen MJ, Frydenberg M, Thastum M, Thomsen PH. The Children's Depression Inventory and classification of major depressive disorder: validity and reliability of the Danish version. *European child & adolescent psychiatry*. 2005 Sep;14(6):328-34.

Sorsdahl AB, Moe-Nilssen R, Strand LI. Observer reliability of the Gross Motor Performance Measure and the Quality of Upper Extremity Skills Test, based on video recordings. *Developmental medicine and child neurology*. 2008 Feb;50(2):146-51.

Spekman NJ, Roth FP. Clinical evaluation of language functions (CELF) diagnostic battery: an analysis and critique. *The Journal of speech and hearing disorders*. 1984 Feb;49(1):97-100.

Spiritos M, O'Mahony P, Malone J. Interrater reliability of the Melbourne Assessment of Unilateral Upper Limb Function for children with hemiplegic cerebral palsy. *Am J Occup Ther*. 2011 Jul-Aug;65(4):378-83.

Spittle AJ, Doyle LW, Boyd RN. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Developmental medicine and child neurology*. 2008 Apr;50(4):254-66.

Sritipsukho P, Wisai M, Thavorncharoensap M. Reliability and validity of the Thai version of the Paediatric Quality of Life Inventory 4.0. *Qual Life Res*. 2012 May 3.

Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held Dynamometry Correlation With the Gold Standard Isokinetic Dynamometry: A Systematic Review. *Pm R*. 2011 May;3(5):472-9.



- Stefanutti D, Fitting JW. Sniff nasal inspiratory pressure. Reference values in Caucasian children. *American journal of respiratory and critical care medicine*. 1999 Jan;159(1):107-11.
- Stewart-Brown S, Patterson J, Mockford C, Barlow J, Klimes I, Pyper C. Impact of a general practice based group parenting programme: quantitative and qualitative results from a controlled trial at 12 months. *Archives of disease in childhood*. 2004 Jun;89(6):519-25.
- Storch A, Jost WH, Vieregge P, Spiegel J, Greulich W, Durner J, *et al*. Randomised, double-blind, placebo-controlled trial on symptomatic effects of coenzyme Q(10) in Parkinson disease. *Arch Neurol*. 2007 Jul;64(7):938-44.
- Stulemeijer M, de Jong LW, Fiselier TJ, Hoogveld SW, Bleijenberg G. Cognitive behaviour therapy for adolescents with chronic fatigue syndrome: randomised controlled trial. *BMJ (Clinical research ed)*. 2005 Jan 1;330(7481):14.
- Subramony SH, May W, Lynch D, Gomez C, Fischbeck K, Hallett M, *et al*. Measuring Friedreich ataxia: Interrater reliability of a neurologic rating scale. *Neurology*. 2005 Apr 12;64(7):1261-2.
- Suomalainen A, Elo JM, Pietilainen KH, Hakonen AH, Sevastianova K, Korpela M, *et al*. FGF-21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. *Lancet Neurol*. 2011 Sep;10(9):806-18.
- Swoboda KJ, Scott CB, Crawford TO, Simard LR, Reyna SP, Krossschell KJ, *et al*. SMA CARNI-VAL trial part I: double-blind, randomised, placebo-controlled trial of L-carnitine and valproic acid in spinal muscular atrophy. *PloS one*. 2010;5(8):e12140.
- Takahashi S, Oki J, Miyamoto A, Okuno A. Proton magnetic resonance spectroscopy to study the metabolic changes in the brain of a patient with Leigh syndrome. *Brain Dev*. 1999 Apr;21(3):200-4.
- Takken T, van der Net J, Engelbert RH, Pater S, Helders PJ. Responsiveness of exercise parameters in children with inflammatory myositis. *Arthritis and rheumatism*. 2008 Jan 15;59(1):59-64.
- Tardieu G, Shentoub S, Delarue R. [Research on a technic for measurement of spasticity]. *Rev Neurol (Paris)*. 1954;91(2):143-4.
- Tartaglia MC, Chen JT, Caramanos Z, Taivassalo T, Arnold DL, Argov Z. Muscle phosphorus magnetic resonance spectroscopy oxidative indices correlate with physical activity. *Muscle & nerve*. 2000 Feb;23(2):175-81.
- Terzi N, Orlikowski D, Fermanian C, Lejaille M, Falaize L, Louis A, *et al*. Measuring inspiratory muscle strength in neuromuscular disease: one test or two? *Eur Respir J*. 2008 Jan;31(1):93-8.
- Thorley M, Lannin N, Cusick A, Novak I, Boyd R. Reliability of the quality of upper extremity skills test for children with cerebral palsy aged 2 to 12 years. *Phys Occup Ther Paediatr*. 2012 Feb;32(1):4-21.
- Timbremont B, Braet C, Dreessen L. Assessing depression in youth: relation between the Children's Depression Inventory and a structured interview. *Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*. 2004 Mar;33(1):149-57.
- Trouillas P, Takayanagi T, Hallett M, Currier RD, Subramony SH, Wessel K, *et al*. International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology. *J Neurol Sci*. 1997 Feb 12;145(2):205-11.
- Tsiros MD, Grimshaw PN, Schield AJ, Buckley JD. Test-retest reliability of the Biodex System 4 Isokinetic Dynamometer for knee strength assessment in paediatric populations. *Journal of allied health*. 2011 Fall;40(3):115-9.
- Tsiros MD, Grimshaw PN, Shield AJ, Buckley JD. The Biodex isokinetic dynamometer for knee strength assessment in children: advantages and limitations. *Work (Reading, Mass)*. 2011;39(2):161-7.

Ullenhag A, Almqvist L, Granlund M, Krumlinde-Sundholm L. Cultural validity of the Children's Assessment of Participation and Enjoyment/Preferences for Activities of Children (CAPE/PAC). *Scand J Occup Ther*. 2011 Dec 16.

Van den Hout JM, Kamphoven JH, Winkel LP, Arts WF, De Klerk JB, Loonen MC, *et al*. Long-term intravenous treatment of Pompe disease with recombinant human alpha-glucosidase from milk. *Paediatrics*. 2004 May;113(5):e448-57.

Van Der Toorn A, Barenbrug P, Snoep G, Van Der Veen FH, Delhaas T, Prinzen FW, *et al*. Transmural gradients of cardiac myofiber shortening in aortic valve stenosis patients using MRI tagging. *American journal of physiology Heart and circulatory physiology*. 2002 Oct;283(4):H1609-15.

van Eldik MCM, Schiliching JEPT, Lutje Spelberg HC, van der Meulen BF, S. vdM. *Reynell Test voor Taalbegrip: Hartcourt Test Publishers*; 2004.

van Hartingsveldt MJ, Cup EH, Oostendorp RA. Reliability and validity of the fine motor scale of the Peabody Developmental Motor Scales-2. *Occup Ther Int*. 2005;12(1):1-13.

van Mater HA, Williams JW, Jr., Coeytaux RR, Sanders GD, Kemper AR. Psychometric characteristics of outcome measures in juvenile idiopathic arthritis: a systematic review. *Arthritis care & research*. 2012 Apr;64(4):554-62.

Vandervelde L, Van den Bergh PY, Goemans N, Thonnard JL. Activity limitations in patients with neuromuscular disorders: a responsiveness study of the ACTIVLIM questionnaire. *Neuromuscul Disord*. 2009 Feb;19(2):99-103.

Vandervelde L, Van den Bergh PY, Goemans N, Thonnard JL. ACTIVLIM: a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscul Disord*. 2007 Jun;17(6):459-69.

Varni JW, Burwinkle TM, Berrin SJ, Sherman SA, Artavia K, Malcarne VL, *et al*. The PedsQL in paediatric cerebral palsy: reliability, validity, and sensitivity of the Generic Core Scales and Cerebral Palsy Module. *Developmental medicine and child neurology*. 2006 Jun;48(6):442-9.

Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in paediatric cancer: reliability and validity of the Paediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer*. 2002 Apr 1;94(7):2090-106.

Varni JW, Burwinkle TM, Limbers CA, Szer IS. The PedsQL as a patient-reported outcome in children and adolescents with fibromyalgia: an analysis of OMERACT domains. *Health and quality of life outcomes*. 2007;5:9.

Varni JW, Burwinkle TM, Seid M. The PedsQL as a paediatric patient-reported outcome: reliability and validity of the PedsQL Measurement Model in 25,000 children. *Expert Rev Pharmacoecon Outcomes Res*. 2005 Dec;5(6):705-19.

Varni JW, Burwinkle TM, Szer IS. The PedsQL Multidimensional Fatigue Scale in paediatric rheumatology: reliability and validity. *J Rheumatol*. 2004 Dec;31(12):2494-500.

Varni JW, Limbers CA. The PedsQL Multidimensional Fatigue Scale in young adults: feasibility, reliability and validity in a University student population. *Qual Life Res*. 2008 Feb;17(1):105-14.

Varni JW, Limbers CA, Bryant WP, Wilson DP. The PedsQL multidimensional fatigue scale in paediatric obesity: feasibility, reliability and validity. *Int J Paediatr Obes*. 2010;5(1):34-42.

Varni JW, Limbers CA, Bryant WP, Wilson DP. The PedsQL Multidimensional Fatigue Scale in type 1 diabetes: feasibility, reliability, and validity. *Paediatric diabetes*. 2009 Aug;10(5):321-8.

Vitale MG, Levy DE, Moskowitz AJ, Gelijs AC, Spellmann M, Verdisco L, *et al*. Capturing quality of life in paediatric orthopaedics: two recent measures compared. *Journal of paediatric orthopedics*. 2001 Sep-Oct;21(5):629-35.

- Voss LD. Can we measure growth? *Journal of medical screening*. 1995;2(3):164-7.
- Vos-Vromans DC, Ketelaar M, Gorter JW. Responsiveness of evaluative measures for children with cerebral palsy: the Gross Motor Function Measure and the Pediatric Evaluation of Disability Inventory. *Disabil Rehabil*. 2005 Oct 30;27(20):1245-52.
- Vuillerot C, Girardot F, Payan C, Fermanian J, Iwaz J, De Lattre C, *et al*. Monitoring changes and predicting loss of ambulation in Duchenne muscular dystrophy with the Motor Function Measure. *Developmental medicine and child neurology*. 2010 Jan;52(1):60-5.
- Wagner JL, Smith G, Ferguson P. Self-efficacy for seizure management and youth depressive symptoms: caregiver and youth perspectives. *Seizure : the journal of the British Epilepsy Association*. 2012 Jun;21(5):334-9.
- Wagner MB, Vignos PJ, Jr., Carlozzi C, Hull AL. Assessment of hand function in Duchenne muscular dystrophy. *Arch Phys Med Rehabil*. 1993 Aug;74(8):801-4.
- Waisbren S, White DA. Screening for cognitive and social-emotional problems in individuals with PKU: tools for use in the metabolic clinic. *Molecular genetics and metabolism*. 2010;99 Suppl 1:S96-9.
- Wallen M, O'Flaherty SJ, Waugh MC. Functional outcomes of intramuscular botulinum toxin type a and occupational therapy in the upper limbs of children with cerebral palsy: a randomised controlled trial. *Arch Phys Med Rehabil*. 2007 Jan;88(1):1-10.
- Wallen MA, O'Flaherty S J, Waugh MC. Functional outcomes of intramuscular botulinum toxin type A in the upper limbs of children with cerebral palsy: a phase II trial. *Arch Phys Med Rehabil*. 2004 Feb;85(2):192-200.
- Warnick EM, Bracken MB, Ksal S. Screening efficiency of the child behavior checklist and strengths and difficulties questionnaire: a systematic review. *Adolesc Ment Health*. 2008(13):140-7.
- Williams EN, Carroll SG, Reddihough DS, Phillips BA, Galea MP. Investigation of the timed 'up & go' test in children. *Developmental medicine and child neurology*. 2005 Aug;47(8):518-24.
- Wraith JE, Clarke LA, Beck M, Kolodny EH, Pastores GM, Muenzer J, *et al*. Enzyme replacement therapy for mucopolysaccharidosis I: a randomised, double-blinded, placebo-controlled, multinational study of recombinant human alpha-L-iduronidase (laronidase). *The Journal of paediatrics*. 2004 May;144(5):581-8.
- Wren TA, Blum S, Tseng-Ong L, Gilsanz V. Three-point technique of fat quantification of muscle tissue as a marker of disease progression in Duchenne muscular dystrophy: preliminary study. *AJR Am J Roentgenol*. 2008 Jan;190(1):W8-12.
- Yoneda M, Maeda M, Kimura H, Fujii A, Katayama K, Kuriyama M. Vasogenic edema on MELAS: a serial study with diffusion-weighted MR imaging. *Neurology*. 1999 Dec 10;53(9):2182-4.
- Young B, Rice H, Dixon-Woods M, Colver AF, Parkinson KN. A qualitative study of the health-related quality of life of disabled children. *Developmental medicine and child neurology*. 2007 Sep;49(9):660-5.
- Young NL, Williams JI, Yoshida KK, Wright JG. Measurement properties of the activities scale for kids. *Journal of clinical epidemiology*. 2000 Feb;53(2):125-37.
- Young NL, Yoshida KK, Williams JI, Bombardier C, Wright JG. The role of children in reporting their physical disability. *Arch Phys Med Rehabil*. 1995 Oct;76(10):913-8.
- Zink I, Lambrechts D. *Nederlandstalige NonSpeech Test (NNST)*: Acco; 2000.
- Zwarts MJ, Bleijenberg G, van Engelen BG. Clinical neurophysiology of fatigue. *Clin Neurophysiol*. 2008 Jan;119(1):2-10.

## Supplementary Document 5

Genetic, biochemical and clinical features of patients with nuclear encoded Complex I deficiency

Online at [http://static-content.springer.com/esm/art%3A10.1007%2Fs10545-012-9492-z/MediaObjects/10545\\_2012\\_9492\\_MOESM1\\_ESM.xls](http://static-content.springer.com/esm/art%3A10.1007%2Fs10545-012-9492-z/MediaObjects/10545_2012_9492_MOESM1_ESM.xls)



## Supplementary Document 6

### Patient characteristics. Genetic, biochemical and clinical details for each patient.

Psychomotor retardation is defined: IQ<50 = severe; IQ 50-70 = mild and IQ >70 = normal.

Clinical characteristics			Genetics		Biochemistry muscle			GMFM	
Patient	Sex	Age (years)	Mutated gene	Heteroplasmy (%)	ATP production (% of lowest reference value)	Complex deficiencies (percentage of lowest reference value)	Clinical phenotype	Myopathy	Encephalopathy
1	F	10	<i>MTTL1</i>	65 (L) / 93 (UEC)	ND	ND	Exercise intolerance	x	
2	F	16	<i>MTTL1</i>	87 (Mu)	67	Normal	Small stature, exercise intolerance	x	x
3	F	16	<i>FRDA</i>		ND	ND	Friedreich ataxia	x	
4	M	16	Pending		50	C III (61%)	Ataxia		x
5	M	11	<i>SDHA</i>		86	C II (26 %) and C II + III (31%)	Dystonia, ataxia, chorea		x
6	F	9	Pending		50	C I, II and III slightly □	Exercise intolerance	x	
7	F	13	<i>AGK</i>		5	CI (26%), CIII (26%), CIV (20%) en CII + CIII (52%)	Sengers-like syndrome	x	
8	M	12	<i>MTTL1</i>	75 (Mu)	100	C I (89%)	Failure to thrive, muscle weakness, exercise intolerance	x	
9#	M	17	<i>RARS2</i>		60	CI (71%) en CIII (83%)	Ataxia	x	x
10#	F	12	Pending		ND	ND	Mild ataxia	x	x
11	F	15	Pending		43	CI (83%) and CIII (82%)	Severe PMR, opticshypoplasia, epilepsy	x	x
12	M	17	Pending		88	CI (86%)	Epilepsy	x	x
13	M	6	<i>MTTS1</i>	97 (L)/96 (UEC)		CI and CIV slightly ↓	Deafness, hyperactivity, behavioural problems	x	x
14	F	11	<i>NGLY1</i>		50	CI (80%) and CII (80%)	Epilepsy, spasticity	x	x
15	M	8	<i>CDH13 + ND1</i>	100	6.2	CI (4%) and CIV (67%)	Hyperactivity, ataxia, spasticity	x	x
16	F	15	<i>FRDA</i>		ND	ND	Friedreich ataxia	x	x
17	M	12	<i>TAZ</i>		38	CI (36%), CIII (55%), II en IV slightly □	Barth syndrome	x	

# siblings; A = Ankle; CI = Complex I; F = female; GMFM = Gross Motor Function Measure; L = leukocytes; M = male; Mu = muscle; ND = not done; PEDI = Pediatric Evaluation of Disability Inventory; PMR = psychomotor retardation; UEC = Urinary epithelial cells

GMFM					GMFM			PEDI			Tardieu
Myopathy	Encephalopathy	PMR	Able to walk	GMFM score (%)	Functional skills (%)			Caregiver assistance (%)			Positive Tardieu
					Self care	Mobility	Social function	Self care	Mobility	Social function	
x		No	Yes	98.0	95	97	95	95	100	100	
x	x	Mild	Yes	99.3	92	100	94	80	100	100	
x		No	No	32.4	81	31	100	65	17	100	
	x	Mild	With help	67.7	70	55	86	58	60	84	A
	x	Severe	No	49.0	58	45	50	60	51	40	
x		No	Yes	100	73	98	97	55	94	100	
x		No	Yes	95.5	95	95	100	95	100	100	
x		No	Yes	100	95	100	100	95	100	100	
x	x	Mild	No	36.5	41	48	83	30	46	48	A
x	x	Mild	Yes	96.3	86	94	95	73	91	76	A
x	x	Severe	With help	24.7	12	6	17	0	2,9	4	
x	x	Mild	Yes	99.6	80	89	86	60	88	64	
x	x	Mild	Yes	96.7	78	98	83	58	100	52	
x	x	Severe	No	8.0	18	6	18	3	0	0	
x	x	Mild	No	44.1	45	54	42	18	60	12	A
x	x	No	No	27.0	70	37	100	33	37	96	
x		No	Yes	99.7	100	97	88	100	100	80	

## Supplementary Document 7

### Correlations between the GMFM and the PEDI and activity parameters measured with the accelerometer.

A level of  $p = 0.0001$  was used for significance; significant  $p$ -values are indicated in bold.

	GMFM	PEDI self-care	PEDI mobility	PEDI - social function
Rest (%)	<b>-0.82</b>	-0.54	<b>-0.87</b>	-0.17
Dynamic activity (%)	0.40	0.25	0.50	0.30
Average counts upper leg (counts/hour)	0.61	0.28	0.70	-0.10
Average counts upper arm (counts/hour)	0.50	0.16	0.49	-0.18
Average counts lower arm (counts/hour)	0.51	0.24	0.60	0.05
Average counts wheelchair (counts/hour)	0.11	0.16	0.09	0.65
Maximal intensity upper leg (counts/min)	0.57	0.18	<b>0.60</b>	-0.07
Maximal intensity upper arm (counts/min)	0.54	0.30	0.48	0.15
Maximal intensity lower arm (counts/min)	<b>0.67</b>	0.26	0.52	0.57
Largest AUC during ½ hour upper leg (counts)	<b>0.65</b>	0.29	<b>0.70</b>	-0.07
Largest AUC during ½ hour upper arm (counts)	0.51	0.96	0.43	-0.22
Largest AUC during ½ hour lower arm (counts)	0.59	0.19	0.56	-0.12

AUC = Area under the curve; GMFM = Gross Motor Function Measure; min = minutes; PEDI = Pediatric Evaluation of Disability Inventory



## Supplementary Document 8

The final version of the International Paediatric Mitochondrial Disease Scale (IPMDS).

Name:

Date of birth:

Date of assessment:

Time since previous IPMDS:

Name physician:

Domain 1:	(raw score)	/(103 - ) =	%
Domain 2:	(raw score)	/(76 - ) =	%
Domain 3:	(raw score)	/(64 - ) =	%
Total score:	(raw score)	/(243 - ) =	%

We kindly ask you to upload your PDF at <http://www.ncmd.nl> to facilitate data collection

**Disease course since previous IPMDS**

Parents:	Stable	Improving	Deteriorating
Physician:	Stable	Improving	Deteriorating

**Over the past 2 days:****1. How is your child's physical condition?**

- A Stable
- B Deteriorating
- C Improving

**2. How is your child's mental health?**

- A Stable
- B Deteriorating
- C Improving

**3. In general, how happy is your child feeling?**

- A Unhappy
- B Not happy nor unhappy
- C Happy
- D Very happy, more happy than others

**4. In general, how comfortable is your child?**

- A Uncomfortable
- B Not comfortable nor uncomfortable
- C Comfortable

## 1. Complaints and symptoms

Please answer the following questions with regard to how your child was complainin over the past 4 weeks:

### 1. Which of the following best describes your child's alertness and response to his/her environment?

- 0 Appropriately alert during daytime hours; good response to the environment
- 1 Alternates between periods of sleepiness and full alertness and/or requires stimulation to remain alert
- 2 Sleepy, only reacting to touch, noise or visual stimulation
- 3 Alternating between periods with no response to environment and periods of sleeping

### 2. How well has your child been able to perform exercise?

- 0 Normal compared to peers
- 1 Limited in playing sports, participating in (adapted) physical education
- 2 Symptomatic on inclines or stairs
- 3 Adaptations to save energy required for playing outside
- 4 Adaptations to save energy required for playing inside
- 5 Restricted on playing inside and frequent resting required despite adaptation of playing inside
- \* Not playing due to severe cognitive impairment

### 3. For how long is your child able to walk?

- 0 Normal compared to peers
- 1 Rest required or impaired coordination after 1 hour of walking
- 2 Rest required or impaired coordination after 30 minutes of walking
- 3 Rest required or impaired coordination after 15 minutes of walking
- 4 Rest required or impaired coordination after 5 minutes of walking
- 5 Not at all

### 4. How tired has your child been?

- 0 Normal
- 1 Tired the day after going on an exciting day out e.g. an amusement park or zoo (with or without use of wheelchair)
- 2 Not able to go on an exciting day out e.g. a amusement park or zoo (with or without use of wheelchair) because of tiredness
- 3 Not able to go on half day family trips e.g. to park or playground (with or without use of wheelchair) because of tiredness
- 4 Not able to go to school for a full week because of tiredness
- 5 Not able to go to school at all because of tiredness

**5. Has your child been depressed, sad, withdrawn, anxious or quiet?**

- 0 Not more than peers
- 1 Only in specific circumstances (tired, exciting events, setbacks, life events)
- 2 Symptoms present every week
- 3 Symptoms present every day but not limiting school performance or the ability to make friends
- 4 Symptoms limiting school performance or the ability to make friends
- 5 Severe depressive symptoms requiring hospitalization
- \* Impossible to indicate due to severe cognitive impairment

**6. Has your child experienced epileptic seizures? At the moment these are treated / untreated \* (\*please circle)**

- 0 No
- 1 Once with self recovery this month (seizure lasted                      sec/min)
- 2 Two or three times with self recovery this month, or absence epilepsy
- 3 One convulsion in which emergency medication was necessary or where there was no recovery within 5 minutes, or more than 3 with self recovery last month
- 4 Multiple or prolonged convulsions in which emergency medication was given or where there was no recovery within 5 minutes
- 5 Hospitalization or more than 3 generalized convulsions (not absences) a week

**7. Has your child suffered from headache?**

- 0 No
- 1 Mild headaches not limiting daily activities
- 2 Daily activities limited by headaches less than once a week
- 3 Daily activities limited at least once a week or true migraine attacks at least once a month
- 4 True migraine attacks despite chronic treatment or true migraine attacks at least once a week
- 5 Chronic migraine: true migraine headaches 15 days/month for longer than 3 months
- \* impossible to indicate

**8. Has your child suffered from muscle pain?**

- 0 No
- 1 Occasional muscle pain after exercise not limiting daily activities
- 2 Recurrent mild muscle pain after exercise not limiting daily activities
- 3 Muscle pain after exercise limiting daily activities
- 4 Spontaneous muscle pain (not only after exercise) limiting daily activities
- 5 Severe, spontaneous muscle pain
- \* impossible to indicate

**9. Has your child suffered any infections?**

- 0 No or only mild infection
- 1 Infection with subsequent tiredness limiting daily activities for more than one day
- 2 Infection with subsequent tiredness limiting daily activities for more than one week
- 3 Infection taking more than one week, with subsequent school absence for more than one day
- 4 Infection taking more than one week, with subsequent school absence for more than one week

**10. Has your child experienced problems with chewing?**

- 0 No
- 1 Bread-crusts and meat are tiring
- 2 Bread-crusts and meat are avoided because of the chewing problems
- 3 Food is pureed because of the chewing problems
- 4 Pureed food is supplemented with tube feeding because of chewing difficulties
- 5 Only tube feeding because of chewing problems

**11. Has your child experienced problems with swallowing?**

- 0 No
- 1 Difficulties (choking) with fluids or dry food
- 2 Difficulties (choking) with all foods; adaptation of the diet
- 3 Choking despite adaptation of the diet (e.g. requiring supplementary tube feeding)
- 4 Solely tube feeding because of choking
- 5 Spontaneous choking on saliva, or recurrent aspiration pneumonia

**12. Has your child experienced hearing problems?**

- 0 No
- 1 Proven hearing loss, without the need for a hearing aid or fully corrected with hearing aid
- 2 Not fully corrected with hearing aid, but no impaired communication
- 3 Not fully corrected with hearing aid, impaired communication
- 4 Reaction to loud sounds (e.g. clapping) only, despite use of a hearing aid
- 5 No reaction to loud sound

**13. Has your child found problems in communicating?**

- 0 Normal, age appropriate communication
- 1 Stammer, dysarthria or language delay impairing communication with strangers
- 2 Stammer, dysarthria or language delay impairing communication with parents
- 3 Effective communication only using alternative methods (speech computer, sign language)
- 4 No effective communication with strangers despite using appropriate alternative methods
- 5 No effective communication with parents despite alternative methods

**14. Has your child been vomiting?**

- 0 No frequent vomiting
- 1 Vomiting at least once a week
- 2 Vomiting once a day despite adaptation of feeding
- 3 Vomiting more than once a day despite adaptation of feeding
- 4 Continuous tube feeding
- 5 Malnourished as a consequence of vomiting despite adaptations in feeding

**15. Has your child been having symptoms of (or medication for) gastroesophageal reflux?**

- 0 No symptoms of gastroesophageal reflux (heartburn, regurgitation, uncomfortable when lying down after consumption of food)
- 1 No symptoms with one medication
- 2 No symptoms with multiple medications
- 3 Symptoms despite multiple medications
- 4 Esophageal erosions despite treatment
- 5 Fundoplication or malnourished despite adaption of feeding
- \* impossible to indicate

**16. Has your child had symptoms of constipation?**

- 0 No constipation
- 1 Variable defecation pattern with defecation between 2x/week and 3x/day
- 2 Constipation fully resolved with oral medication
- 3 Constipation not fully resolved with oral medication (frequency 1-2x per week)
- 4 Constipation with frequency <1x/week despite oral medication
- 5 Constipation requiring clysters or colonic lavage

**17. Has your child had symptoms of diarrhoea?**

- 0 No diarrhoea
- 1 Loose stool  $\geq$  3x/day
- 2 Shapeless or watery stool < 3x/day
- 3 Shapeless or watery stool  $\geq$  3x/day
- 4 Episodes of diarrhoea that require a change of clothes, occurring at least once per week
- 5 Severe diarrhoea necessitating increased fluid supply

**18. Cognitive development**

- 0 Normal cognitive development, mainstream school or attending special school only because of physical disability
- 1 Attending special school because of mental disability but learning new skills
- 2 Attending special school because of mental disability but not learning new skills
- 3 Losing skills in one area (cognitive, language, motor, social/emotional)
- 4 Losing skills in more than one area or severe cognitive impairment due to prior loss of skills

**19. Does your child have behavioural problems?**

- 0 Normal compared to peers
- 1 More severe behavioural problems compared to peers
- 2 Behavioural problems limiting the ability of the family to go out
- 3 Behavioural problems limiting the ability to make friends
- 4 Behavioural problems limiting the ability to go to school
- 5 Not attending school because of behavioural problems

**20. Does your child have autistic features?**

- 0 Normal compared to peers
  - 1 No official diagnosis of an autism spectrum disorder. Parents do recognize autistic features which cannot be explained by mental retardation, in their child
  - 2 Official diagnosis of an autism spectrum disorder. Limited social interaction, limited flexibility of behavior or difficulty switching between activities requiring support, but not limiting daily functioning
  - 3 Official diagnosis of an autism spectrum disorder. Clear problems in social interaction, difficulty with change and limited flexibility in behavior, limiting daily functioning
  - 4 Official diagnosis of an autism spectrum disorder. Severe deficit in social communication, great difficulty with change and very limited flexibility in behavior despite the use of medication
- \* impossible to indicate due to severe cognitive impairment

**21. Breathing pattern**

- 0 Normal breathing pattern
- 1 Frequent sighing or irregular breathing
- 2 Nocturnal ventilator support required
- 3 Periods of apnoea (> 20 sec) or continuous ventilation

**22. Has your child been continent of urine (able to indicate when toilet visit is needed > 90% of time)?**

- |                  |                 |
|------------------|-----------------|
| During the day   | Yes (0) /No (1) |
| During the night | Yes (0) /No (1) |

**23. Does your child experience any of the following items when tired?**

- |                              |                 |
|------------------------------|-----------------|
| Strabismus (squint)          | Yes (1) /No (0) |
| Ptosis (drooping eyelids)    | Yes (1) /No (0) |
| Dysarthria (impaired speech) | Yes (1) /No (0) |

## 2. Physical examination

In the case of asymmetrical abnormalities, score the most severely affected side

### 1. Growth

- 0 Normal (-2 to +2 SD) for target height/catch up growth
- 1 Growing parallel to growth curve, but below -2SD for target height
- 2 Deviating from his/her own growth curve appropriate for target height
- 3 Deviating from his/her own growth curve, and below -2SD for target height

### 2. Weight for height

- 0 Normal or catch up growth when under -2 SD
- 1 Growing parallel to -2 SD or deviating ( $\geq 1$  SD) in positive direction in children above +2 SD
- 2 Deflecting ( $\geq 1$  SD) from growth curve in negative direction
- 3 Unintentionally losing  $\geq 3\%$  weight compared to the last measurement

### 3. Alertness

- 0 Eyes opened spontaneously
- 1 Opens eyes when talking (verbal child) or to parent's voice
- 3 Opens eyes when touched
- 3 Opens eyes on pain stimulus
- 4 Not opening eyes

### 4. Breathing pattern

- 0 Normal breathing pattern
- 1 Sighing or irregular breathing observed during physical examination
- 2 Periods of apnoea (> 20 sec) during examination

### 5. Dysarthria

- 0 No dysarthria
- 1 Nasal, hoarse, low pitch voice, or irregular speech, but easily understood
- 2 Obvious dysarthria, understood more than 50% of the time
- 3 Obvious dysarthria, understood less than 50% of the time
- 4 Not able to communicate verbally or use of alternative methods primarily due to dysarthria
- \* Not able to communicate due to other causes



**6. Ptosis**

0 Normal

1 Mild ptosis not obscuring either pupil

2 Unilateral ptosis obscuring &gt; 1/3 of pupil

3 Bilateral ptosis obscuring &gt; 1/3 of pupils

4 Bilateral ptosis obscuring &gt; 2/3 of pupils, or previous ptosis surgery

**7. Strabismus**

0 Normal

1 Permanent strabismus of one eye

2 Permanent strabismus of both eyes

**8. Eye movements**

0 Normal

1 Some restriction in any of the eye movements with normal abduction

2 Partial abduction possible

3 Abduction minimal

**9. Nystagmus**

0 No or physiological (endpoint) nystagmus

1 Gaze evoked nystagmus

2 Spontaneous nystagmus

**10. Vision – using glasses and both eyes**0  $\geq 1.0$  (Snellen or crowded LogMAR chart) or attention to small objects at a distance1  $\geq 0.5 - < 1.0$  (Snellen or crowded LogMAR chart)2  $\geq 0.1 - < 0.5$  (Snellen or crowded LogMAR chart) 3  $< 0.1$  (Snellen or crowded LogMAR chart)

4 Only reaction to big, colored objects within reach

5 No reaction to big, coloured objects within visual field

\* Impossible to indicate due to severe cognitive impairment

**11. Proximal muscle strength**

0 Normal power in arms and legs

1 Moves legs and arms against resistance, but not against full resistance

2 Moves arms and legs against gravity, but not against resistance

3 Active joint movement when gravity is eliminated (e.g. in horizontal plane)

4 Contraction of muscle is visible, but no movement of the joint

5 No contraction visible

## **12. Distal muscle strength**

- 0 Normal power in feet and hands
- 1 Moves hands and feet against resistance, but not against full resistance
- 2 Moves hands and feet against gravity, but not against resistance
- 3 Active joint movement when gravity is eliminated (e.g. on a underground)
- 4 Contraction of muscle is visible, but no movement of the joint
- 5 No contraction visible

## **13. Hypokinesia**

- 0 Absent or reduced facial expression due to other causes (e.g. facies myopathica)
- 1 Reduced facial expression only, no slowing of body movements
- 2 Reduced facial expression, mild slowing of body movements
- 3 Significant slowing of body movements

## **14. Abnormal, involuntary movements and abnormal posturing**

- 0 Absent
- 1 Intermittent involuntary movements of one extremity, not interfering with daily activities
- 2 Intermittent involuntary movements of multiple extremities, mildly interfering with daily activities
- 3 Continuous involuntary movements of whole body, markedly interfering with daily activities

## **15. Ataxia**

- 0 Absent
- 1 Impaired coordination (dysmetria) compared to peers on examination only (hesitant heel-toe, disturbed alternate movements)
- 2 Gait reasonably steady, but unable to maintain heel-toe, mild dysmetria, truncal ataxia, not limiting the ability to sit
- 3 Ataxic gait, not able to walk 2 passes heel-toe, past pointing with intention tremor, truncal ataxia, limiting the ability to sit
- 4 Walking unsafe or sitting impossible without support, primarily due to ataxia

## **16. Tremor**

- 0 Absent or physiological tremor
- 1 Tremor present at rest, but not affecting coordination
- 2 Tremor affecting coordination
- 3 Unable to grab things primarily due to tremor

**17. Reflexes**

- 0 Normal, brisk, indifferent or hypoactive reflexes
- 1 Abnormally brisk reflexes
- 2 Expansion of the reflex zone or crossed adductor response
- 3 Non sustained clonus
- 4 Sustained clonus

**18. Hypertonia**

- 0 Normal tone or hypotonic
- 1 Increased tone, catch followed by relaxation or minimal resistance at the end of the range of motion
- 2 Increased tone, catch followed by minimal resistance during less than the half of the range of motion
- 3 Resistance during most of the whole range of motion
- 4 Resistance, passive movements are difficult
- 5 Rigid flexion or extension

**19. Hypotonia**

- 0 Normal tone or hypertonic
- 1 Mild slipping through/head lag or hypotonia or hypermobility on physical examination
- 2 Obvious slipping through/head lag, sits with curved back
- 3 Obvious slipping through/head lag, difficulty maintaining the head in midline position while sitting with lower backsupport
- 4 Obvious slipping through/head lag, unable to sit with or without backsupport only primarily due to hypotonia

**20. Rigidity**

- 0 No rigidity or hypertonia only
- 1 Paratonia
- 2 Mild and inconsistent resistance throughout the range of motion
- 3 Resistance is detected consistently during the range of motion, but full range of motion is obtained easily
- 4 Resistance to obtain passive movement requires maximal effort by the rater or full range of motion is not obtained primarily due to rigidity

**21. Sensory examination**

- |              |  |
|--------------|--|
| Vibration    | Normal (0) / Abnormal (1) / Absent (2)/Impossible to indicate (*)  |
| Subtle touch | Normal (0) / Abnormal (1) / Absent (2) /Impossible to indicate (*) |

### 3. Functional tests

Some, indicated items are only for children  $\geq 6$  years

#### 1. Communication

- 0 Actively interacting with researcher, easily understood, follows complex instructions
- 1 Actively interacting with researcher; difficulty expressing self; uses alternative methods to communicate; follows simple instructions
- 2 Actively interacting with researcher; difficult to understand despite use of alternative communication methods; child has difficulty understanding simple instructions
- 3 Reactive to researcher but no understanding of simple single-level instruction and not understood by researcher
- 4 Reacts only to tactile, auditory or visual stimulation
- 5 No reaction to tactile, auditory or visual stimulation

#### 2. Head control – maintains head in midline while sitting with back support

- 0 Normal
- 1 Control, but only with back support
- 2 Control longer than 1 min
- 3 Control between 30 seconds and 1 min
- 4 Control shorter than 30 seconds
- 5 Absent

#### 3. Rolling over – from supine to prone and back

- 0 Able to roll over from supine to prone and back over both sides
- 1 Able to roll over from supine to prone and back, but only to left or right
- 2 Able to roll over from supine to prone and back, but only to supine or prone
- 3 Initiates movement and lifts shoulder and hip from underground
- 4 Initiates movement and moves arm over the midline
- 5 No attempt to roll over

#### 4. Sitting up – from lying supine to sitting

- 0 Able to sit up without help of elbows or hands
- 1 Able to sit up with help of elbows or hands
- 2 Able to sit up using a trick manoeuvre (e.g. turns to prone side)
- 3 Lifts head from ground, but unable to sit up
- 4 Not able to lift head from the ground
- 5 No attempt to sit up

**5. Sitting position – sits without back support**

0 Normal

- 1 Able to sit without support for at least 5 seconds, but not able to hold arms upright for 5 seconds
- 2 Able to sit without support, but for less than 5 seconds
- 3 Able to sit only with support of the hands
- 4 Not able to sit

**6. Standing up from a chair [chair at a height that allows both feet to be resting flat on the ground] if possible without support of hands**

0 Normal

- 1 Compensatory movements, but within normal time and no help of hands
- 2 Slips off, needs own hands for support or abnormal method, but within normal time
- 3 Takes longer than expected for age, due to cognitive or motor inabilities
- 4 Only possible with help of caretaker, or unsafe procedure
- 5 Not able to stand up

**7. Standing – stands without support**

0 Normal

- 1 Compensatory movements or abnormal posture
- 2 Stands for less than one minute without support
- 3 Stands for less than 30 seconds without support
- 4 Stands (>10 seconds) only with support
- 5 No attempt

**8. Walking – walking without support**

0 Normal

- 1 Compensatory movements or abnormal posture
- 2 Walks for more than 10 m, but not for 20 m
- 3 Walks for less than 10 m
- 4 Walks only a few (1-3 steps) or only walks with support
- 5 No attempt to walk or not able to walk with support

**9. Only ≥ 6years: Running**

0 Normal

- 1 Normal, but slower
- 2 Compensatory movements or abnormal posture
- 3 Decreased time of lifting both feet from the floor (levitation)
- 4 Increases speed but no levitation (moment that both feet are lifted from the floor)
- 5 Absent

**10. Only  $\geq$  6 years: Hopping on one foot – 10 times**

- 0 Normal
- 1 Normal but decreasing hopping height during exercise; OR able to hop normally more than 5 times
- 2 Hops 10 times, starts with toes from the floor, finishes not lifting toes from the floor; OR able to hop normally more than 3 times
- 3 Hops 10 times but not lifting toes from the floor
- 4 Stands on one foot
- 5 No attempt

**11. Reaching – grasp pen/small toy held in the air by the examiner within arm length of the patient**

- 0 Normal
- 1 Abnormal movements
- 2 Abnormal movements, much slower to reach object
- 3 Possible to approach the item, but not grab it
- 4 Only inefficient/uncoordinated efforts not approaching the item
- 5 No attempt

**12. Grasping – grasps pen within arm length from the table and transfers pen/small toy to other hand**

- 0 Normal
- 1 Abnormal movements, but able to achieve both grasping and transferring
- 2 Able to grasp but not transfer
- 3 Able to approach the item, but not grasp it
- 4 Only inefficient/uncoordinated efforts not approaching the item
- 5 No attempt

**13. Only  $\geq$  6 years: rotates pen within the hand with fingers**

- 0 Normal
- 1 Rotates the pen, but with compensatory movements or tricks
- 2 Initiates rotating movement, but unable to turn pen around
- 3 Opens hand with pen in it, but unable to initiate rotating movement
- 4 Attempts, but fails to open hand
- 5 No attempt

## Supplementary Document 9

The final version of the manual for the International Paediatric Mitochondrial Disease Scale (IPMDS).

### Introduction

The IPMDS was designed to monitor general disease progression of symptoms associated with mitochondrial disease in children (aged 0 – 18 years) with a mitochondrial disorder. The scale provides detailed insight into the complaints and symptoms (domain 1), physical examination (domain 2) and functioning of the patient (domain 3).

### PREREQUISITS

#### TRAINING

In order to familiarize examiners with the utilization of the IPMDS and its scoring principles, a training session, in which the items are practiced in a role-play at least twice, is recommended.

Please also watch our instructional video.

#### MATERIAL

- A complete growth chart of the child
- A Snellen or crowded LogMAR chart at least 3 m from a wall
- A physiotherapy mat or a wide examination table
- A chair (with adjustable height if possible). The height of the seat should be such that the person's feet touch the floor or footstool when seated (with or without footstool)
- A footstool if indicated
- A pillow for back support if indicated
- Space within a corridor, wide at least 2m, to enable free movement of 20 m, with marks at 10 and 20 m
- A marked line traced on the floor, approximately 3 m long and 2 cm wide.
- A penlight
- A reflex hammer
- A tuning fork
- A pen

## **PATIENT**

The patient should be wearing comfortable clothes, which will not interfere with movements. Corsets should be removed, if possible. Patients should be barefooted without orthoses and socks.

## **SPECIFIC INSTRUCTION**

The order of the items should be respected.

## **DOMAIN 1**

The questions in domain 1 should be assessed with a caregiver/patient interview and not by chart review nor the physician's own judgment.

All questions enquire about the severity of the complaints and symptoms over the past 4 weeks.

The answer "normal" should always be verified by asking for the symptoms rated 1 or 2 points (e.g in question 1 "So he doesn't need stimulation to remain alert? And there are no periods of sleepiness during the day?").

## **DOMAIN 2**

The items in domain 2 are assessed by physical examination. In the case of asymmetrical abnormalities, the worse side is scored. If the item is difficult to interpret, please note this on the scoring sheet.

### Item specific instructions

#### *Item 1: Growth*

The height curve of the past 6 months is used to determine this item.

The individual target height of the child is calculated as following

- o Girls:  $(\text{height father} + \text{height mother} - 13) / 2 + 4.5$  ( in cm)
- o Boys:  $(\text{height father} + \text{height mother} + 13) / 2 + 4.5$  ( in cm)

#### *Item 2: Weight for height*

The weight for height curve of the past 6 months is used to determine this item.

#### *Item 3: Alertness*

No specific instructions, may be observed during the completion of domain 1



*Item 4: Breathing pattern*

No specific instructions, sighing and breathing pattern may be observed during the completion of domain 1

*Item 5: Dysarthria*

May be observed during the completion of domain 1. The type of dysarthria (spastic, myopathic, extrapyramidal, ataxic) is not relevant, the consequences for communication are. If the child is not able to communicate for other reasons, score a \*.

*Item 6: Ptosis*

Ptosis is observed when the child is looking forward with his/her head in a neutral position. Ask for previous ptosis surgery.

*Item 7: Strabismus*

The light reflex in both eyes is observed when looking forward. In case of strabismus, the light reflex will be asymmetrical (not on the pupil).

*Item 8: Eye movement*

The child is asked or stimulated to follow a penlight or an interesting object from left to right; up and down; and up and down when the eyes are deviated left and right. Especially the abduction (pull out) of both eyes is observed.

*Item 9: Nystagmus*

Nystagmus is observed by looking at the blood vessels on the sclera or the iris during the eye movement examination. Physiological nystagmus (small-amplitude ( $<2^\circ$ ) conjugate jerk nystagmus on far eccentric gaze ( $>40^\circ$ )) is rated as a 0. When nystagmus is observed when the child's looking forward, this is noted as spontaneous nystagmus.

*Item 10: Vision*

The child sits or stands at a known distance (e.g. 3m) from the Snellen chart with both eyes opened and using usual glasses. Visual acuity is calculated as (distance to chart (e.g. 3m) / number behind last row adequately seen). If the child is not able to read, the circle (O) can be searched for ("Is this a circle?") or the response to visual stimuli is noted. In case of cognitive impairment, note the reaction to "interesting objects" at a distance.

*Item 11: Proximal muscle strength*

The m. biceps brachii and the m. quadriceps femoris are tested bilaterally. The strength of the child is compared to peers when differentiating between a score of 0 and 1. Gravity is eliminated by asking to move the joint in a horizontal plane on a surface.

*Item 12: Distal muscle strength*

The grip strength and the m. tibialis anterior are tested bilaterally. The strength of the child is compared to peers when differentiating between a score of 0 and 1. Gravity is eliminated by asking to move the joint in a horizontal plane on a surface.

*Item 13: Hypokinesia*

The facial expressions are observed during interaction. Decreased facial expression due to other causes (e.g. facies myopathica) is not rated in this item. Slowing of body movements (hypokinesia) is observed during the other items of the physical examination, walking or playing.

*Item 14: Abnormal, involuntary movements and abnormal posturing*

The presence of dystonia (abnormal positioning of e.g. hands), chorea (quick, jerky movements), athetosis (slow, writhing movements) or hemiballismus (unilateral, flinging movements) is assessed. The effect of these movement disorders on coordination is observed during (un)dressing, walking or playing.

Tics and tremor are not scored at this item.

*Item 15: Ataxia*

Ataxia is tested by heel-toe, alternate movements, the tap nose test and – if abnormal - the ability to sit. The performance on the heel-toe and alternate movements should be compared to peers. Abnormal coordination without ataxia is not rated in this item.

*Item 16: Tremor*

The presence of a tremor is tested by asking the child to put both hands (palm down) forward with straight arms. A physiological tremor is defined as a tremor of 10 Hz.

*Item 17: Reflexes*

Deep tendon reflexes (patellar reflexes, ankle jerk reflexes and biceps brachii reflexes) including reflex spread are tested bilaterally. Physiologically brisk but normal reflexes are rated as a 0. Expansion of the reflex zone is reflex contraction after tapping on the tibia instead of the knee tendon. The crossed conductor response is reflex contraction of the right muscle when tapping on the left knee tendon (and vice versa). The worse side is scored.

*Item 18: Hypertonia*

The presence of hypertonia is assessed by flexion and full extension of the ankles, knees and elbows bilaterally, at slow and at fast speed when the child is relaxed or distracted. The presence of spastic resistance or a catch is observed. Contractures are looked for by fully extending and flexing the ankles, knees and elbows.

*Item 19: Hypotonia*

Muscle tone is assessed by holding the child under the axillae and by placing the child in sitting position. In older children, the tone is assessed during passive movement of the elbow and observing the sitting position.

*Item 20: Rigidity*

The presence of hypertonia is assessed by flexion and full extension of the ankles, knees and elbows bilaterally, at slow speed when the child is relaxed or distracted. The presence of (cogwheel) rigidity is rated. Paratonia, the active resistance of the patient to any movement despite distraction, often augmented by increasing the speed of the movement and often interpreted as “poor cooperation”, is rated as a 1. When the full range of motion cannot be obtained due to contractures or spasticity, the rigidity over the remaining free trajectory is rated.

*Item 21: Sensory examination*

Vibration and subtle touch are tested on the big toe of both feet. The child is asked to close his eyes and say if he can feel that he's being touched, in children not able to respond appropriately, the reaction to the stimulus is noted.

**DOMAIN 3**

In domain 3, the functional abilities of the child are assessed. For this domain, the reason why the child is not able to sit (PMR, ataxia, hypotonia, contractures) is not relevant.

Three attempts are allowed if the investigator expects the child to be able to perform better. Using tricks during three attempts is rated as “compensatory movements”. If the child refuses to attempt the item, if the item was forgotten, if the item is not safe to execute or if the child is not able to maintain the starting position, the score for the item is 5. “Refusal” or “forgotten” should be noted on the scoring sheet.

Some items are dependent on the normal development of the child and are therefore only applicable to patients aged 6 years and older. This is indicated at the item (item 10 and 13).

Investigators are allowed to give an example of the intended movement at every attempt. The instructor is allowed to show the movement by moving the child's limb in the intended direction (e.g. by letting the pen rotate in his/her hand and asking the child to continue).

### Item specific instructions

#### *Item 1. Communication*

As observed during the assessment of domain 1 and 2. Complex instructions are double instructions (do this followed by that) such as the verbal instructions for the dysdiadochokinesis and ataxia exercises. Simple instructions include single instructions (get on the examination table). The cause of the difficulty expressing him/herself (PMR, autism, dysarthria) is not relevant.

#### *Item 2. Head control*

Adequate back support is a straight back of a (wheel)chair, supporting at least ¾ of the back, but not more than shoulder height. The researcher is allowed to position the head above the pelvis, if necessary; the ability to remain the head in this position is scored.

#### *Item 3. Rolling over*

Child is lying flat on his/her back with the head in the midline on a broad examination table or on a hard mat on the ground. The child is asked or tempted (with an interesting object) to roll over to both sides and back.

#### *Item 4. Sitting up*

Child is lying flat on his/her back on a hard mat on the ground. The child is asked or encouraged (with an interesting object) to sit up, if possible without hands (e.g. by asking the child to stab the arms forward).

#### *Item 5. Sitting position*

The child sits without backsupport on a hard mat on the ground. The child is asked to sit without support for 5 seconds and then hold their arms upright for 5 seconds.

#### *Item 6. Standing up from a chair*

The child sits at a chair with a hard surface at a height that allows both feet to be resting flat on the ground or on a footstool. The hands are resting on the lap of the child. The child is asked to raise from the chair without using the hand (e.g. by placing them on their head).

#### *Item 7. Standing*

The child stands on a hard surface in a room or corridor of at least 2m wide. The researcher is behind of the child to catch him/her if he falls, but so that the child cannot seek physical support.

The child is asked to stand straight for one minute.

*Item 8. Walking*

The child stands on a hard surface in a room or corridor of at least 2m wide. The researcher is behind of the child to catch him if he falls but so that the child cannot seek physical support.

The child is asked to walk for 20 m at his/her usual pace.

*Item 9. Only  $\geq 6$  years: Running*

The child stands on a hard surface in a room or corridor of at least 2m wide. The researcher is behind of the child to catch him if he falls but so that the child cannot seek physical support.

The child is asked to run or walk as fast as possible for 20 m.

*Item 10. Only  $\geq 6$  years: Hopping on one foot*

The child stands on a hard surface in a room or corridor of at least 2m wide. The researcher is behind of the child to catch him if he falls but so that the child cannot seek physical support.

The child is asked to hop on one leg for 10 times. One hop is take off and landing at the same single foot.

*Item 11. Reaching*

The child sits stably in a (wheel)chair with adequate back support at a height that allows both feet to be resting flat on the ground or on a footstool. The table is at the height of the elbow or maximally halfway the height of the elbow and the shoulder in resting position. The child is asked or encouraged to reach for a pen or a toy of approximately similar size and weight. The pen/toy is held in the air within arms length of the child.

*Item 12. Grasping*

The child sits stably in a (wheel)chair with adequate back support at a height that allows both feet to be resting flat on the ground or on a footstool. The table is at the height of the elbow or maximally halfway the height of the elbow and the shoulder in resting position. The child is asked or encouraged to grasp a pen or a toy of approximately similar size and weight from the table and to put it in his/her other hand and back. The pen is placed within arm length of the child.

*Item 13. Only  $\geq 6$  years: Rotating*

The child sits stably in a (wheel)chair with adequate back support at a height that allows both feet to be resting flat on the ground or on a footstool. The table is at the height of the elbow or maximally halfway the height of the elbow and the shoulder in resting position.

Rotating the pen is turning the pen around in the hand like a majorette (see instructional video).

## **INTERPRETATION**

The score is expressed as the percentage of items which were feasible to perform. Asterixes (\*) can be scored as well, the total score will change accordingly. E.g. if the parents are not able to indicate the presence of headache, the maximum score of the first domain changes from 103 to 73. If the child is not cooperative during the execution of domain 2 and 3, these items are omitted from the total score.

## Supplementary Document 10

Link to the instructional video.

<https://youtu.be/WxTY-jvbFtg>

## Supplementary Document 11

Inter- and intra-rater and test-retest reliability for the “pilot scoring list – final version”.

The pilot scoring list was adapted based on the experiences of the physicians and the inter- and intra-rater reliability (ICCagreement  $\geq 0.7$  was used as ‘acceptable’; in green). Adaptations of items are indicated in the “items”-column.

	Item	Inter-rater	Test-retest	Intra-rater
S1	Response to environment	NV	NV	NV
S2	Energy	0.798	0.873	1.000
S3	Exercise tolerance	0.393	0.821	0.967
S4	Tired	0.347	0.712	1.000
S5	Depressed	0.921	0.391	1.000
S6	Epilepsy	1.000	1.000	1.000
S7	Headache	0.956	0.000	0.951
S8	Muscle pain	0.883	0.000	0.975
S9	Infections	0.271	1.00	1.000
S10	Chewing	0.837	0.632	1.000
S11	Swallowing	0.945	0.889	1.000
S12	Hearing	0.789	0.848	0.926
S13	Communicating	0.884	0.831	0.761
S14	Vomiting	0.642	NV	NV
S15	GE reflux	0.642	0.774	0.873
S16	Constipation	0.974	0.774	1.000
S17	Diarrhoea	0.761	NV	NV
S18.1	Continence (day)	0.761	0.774	1.000
S18.2	Continence (night)	0.761	1.000	1.000
S19.1	Aggression	0.761	1.000	1.000
S19.2	Routine	0.517	1.000	0.774

S19.3	Fixation themes/objects	0.602	0.533	1.000
S19.4	Sleep pattern	0.364	1.000	0.741
S19.5	Imaginative play	NV	0.741	NV
PE1	Growth	0.671		NV
PE2	Weight	0.046		NV
PE3	Development	0.739		0.906
PE4	Alertness	NV		NV
PE5	Breathing pattern	0.834		1.000
PE6	Dysarthria	0.893		0.979
PE7	Eye movement	NV		NV
PE8	Strabismus	0.460		0.437
PE9	Ptosis	0.342		1.000
PE10	Vision	0.750		0.889
PE11	Hypotonia	0.102		
PE12	Muscle strength (proximal)	0.328		
PE13	Muscle strength (distal)	0.375		
PE14	Hypokinesia	0.518		0.911
PE15	Abnormal movements	0.382		0.926
PE16	Ataxia	0.569		0.962
PE17	Tremor	0.030		1.000
PE18	Pyramidal signs	0.071		
PE19	Hypertonia	NV		
PE20.1	Reflexes	0.125		NV
PE20.2	Hyperreflexia	0.000		NV
PE21.1	Vibration	0.186		NV
PE21.2	Subtle touch	NV		NV
PE21.3	Cold feet	NV		
PE21.4	Cold shins	0.000		
FT1	Communication	0.800		1.000
FT2	Head control	0.000		NV
FT3	Rolling over	0.913		NV
FT4	Sitting up	0.250		1.000
FT5	Sitting	NV		0.710
FT6	Standing up	0.896		1.000
FT7	Standing	0.967		0.951
FT8	Walking	0.965		0.979
FT9	Running	0.825		0.784



FT10	Hopping	0.883		0.916
FT11	Reaching	0.325		0.625
FT12	Grasping	0.365		0.625
FT13	Rotating	0.850		0.962

	minor revision		ICC agreement $\geq 0.70$	
	major revision		ICC agreement $\leq 0.30$	
	removed		NV = no variance	

Main clinical features	Feasibility total	Feasibility of applicable items	NPWDS-S1	NPWDS-S2	NPWDS-S3	NPWDS-total	PEDI selfcare	PEDI mobility
Microcephaly, dysmorphisms, failure to thrive, tired, exercise intolerance	85	96	6	0	10	16	49	55
Pyramidal syndrome, dystonia, rigidity	81	93	11	1	15	27	24	21
Psychiatric, dyskinetic	100	100	8	0	10	18	70	54
Exercise intolerance	100	100	2	5	2	9	73	59
OPEO, deafness, short stature, retinitis pigmentosa	98	98	6	3	10	19	73	59
Ataxia, dystonia, cognitive impairment, dysphagia	95	96	9	0	17	26	52	35
Ataxia, dystonia, cognitive impairment, dysphagia, dysarthria	96	97	16	0	19	35	13	1
Cerebellar ataxia, cortical visual impairment, learning disabilities	98	98	5	5	9	19	70	49
Dystonia, cortical visual impairment, oromotor dysfunction, severe global developmental delay	80	91	0	2	1	3	71	57
Myopathy, dystonia, neuropathy, respiratory failure, oromotor dysfunction, learning difficulties	85	86	18	2	14	34	0	0
Cerebellar ataxia	87	96	16	5	22	43	5	0
Severe developmental delay, severe failure to thrive, visual impairment, intermittent dystonia, epilepsy	91	97	14	3	15	32	4	2
An episode resembling ADEM, later basal ganglia infarcts, neuropathy, epilepsy	97	98	6	4	9	19	73	59
Cognitive regression, intractable epilepsy, visual impairment, mood swing	91	99	0	0	6	6	73	58
Deafness and ataxia	99	99	9	1	5	15	67	41
EPC, acute liver failure, FTT	84	93	14	6	9	29	2	3
Regression, infantile spasms, FTT	84	96	9	9	8	26	2	3



Supplementary Document 12

Patient characteristics for the international field testing.

Patient characteristics, including biochemical and – if known – genetic diagnosis. The IPMDS total and sub domain scores are given, as well as subjective disease severity as rated by parents and by the physicians (mean), the NPMDS total score and sub domain scores and the PEDI total and sub domain scores.

Centre	Patient	Gender	Age (years)	Ethnicity	Age at diagnosis	Syndrome	Biochemical defect	Genetic defect
R	1	F	7	Dutch	2	Leigh syndrome	Complex V deficiency	<i>m.9155A&gt;G</i>
R	2	F	11	Dutch	8	Leigh syndrome	Complex I deficiency	<i>NDUFS7</i>
R	3	M	13	Dutch	3	Leigh syndrome		<i>MTFMT</i>
R	4	M	14	Dutch	12	Alper's syndrome		<i>POLG</i>
Mu	1	M	13	German	10	Kearns Sayre syndrome		<i>single mtDNA deletion</i>
Mu	2	F	16	German	2.5	Leigh syndrome	Complex IV deficiency	<i>SURF1</i>
Mu	3	F	14	German	0.7	Leigh syndrome	Complex IV deficiency	<i>SURF1</i>
HK	1	M	16	Chinese	11	MELAS	Complex I deficiency	<i>m.13513G&gt;A</i>
HK	2	F	3	Chinese	2.8		Complex II and III deficiency (CoQ10 deficiency)	Pending
HK	3	M	9	Chinese	2.9	Leigh syndrome	Complex IV deficiency	<i>SURF1</i>
HK	4	F	8	Chinese	7		Complex IV deficiency	Pending
P	1	M	2	Asian	2		Complex I deficiency	Pending
P	2	M	12	African	6		Complex III deficiency	Pending
P	3	M	12	European	10		Complex I deficiency	Pending
P	4	M	15	European	6		Complex III deficiency	Pending
C	1	F	2.2	Caucasian	1	Alper's syndrome	N/A	<i>POLG</i>
C	2	F	1.6	Caucasian	0.7	Leigh syndrome	N/A	<i>NDUFS8</i>



PEDI social	PEDI ass selfcare	PEDI ass mobility	PEDI ass social	Parents - general disease severity	Parents - functional limitations	General disease severity (mean)	Mito disease severity (mean)	Abnorma- lities at PE (mean)	Severity of functional limitations (mean)	Mild-mode- rate-severe (median)
60	29	32	23	7	7	7	7	7	7	2
65	20	9		8	8	9	8	8	9	3
52	40	35	25	6	7	5	6	4	6	2
65	40	35	25	2	2	4	4	2	3	1
65	40	35	24	7	6	5	7	4	4	2
55	19	19	17	3	3	8	7	7	8	3
52	2	0	10	4	5	9	9	9	9	3
56	37	27	15	6	8	5	5	4	4	2
60	38	32	24	0	0	1	1	1	0	1
0	0	0	0	9	10	9	9	9	10	3
13	1	0	2	10	10	10	10	10	10	3
2	0	0	0	8	10	10	10	10	10	3
65	40	35	25	6	7	9	9	9	10	3
65	40	35	25	4	3	5	5	5	5	2
65	35	18	25	8	8	8	7	8	8	3
3	0	0	0	7	9	8	8	8	8	3
3	0	0	0	8	9	8	7	7	8	3



ADEM = acute disseminated encephalomyelitis; C = Childrens' Hospital of Philadelphia; CPEO = chronic progressive external ophthalmoplegia; EPO = epilepsy partialis continua; F = female; FTT = failure to thrive; HK = the University of Hong Kong; M = male; MELAS = mitochondrial encephalopathy with lactic acidosis and stroke-like episodes; MU = Ludwig Maximilians Universität München; N/A = not assessed; NPMDS = Newcastle Paediatric Mitochondrial Disease Scale; P = University of Pretoria; PEDI = Pediatric Evaluation of Disability Inventory; R = ErasmusMC Rotterdam

## Supplementary Document 13

### Rater characteristics for the international field testing.

Centre	Rater	Year of graduation from medical school	Highest education	Years of experience in mitochondrial field	Time spent with mitochondrial patients per month (hours)
R	1	2003	Paediatric Neurology	0	1
R	2	2003	Neurology	0.5	0
R	3	1981	Neurology	25	16
R	4	2005	Neurology	0	0
Mu	1	2005	Neurology	1	32
Mu	2	2005	Neurology	5	32
Mu	3	2003	Paediatric Neurology	11	46
Mu	4	1990	Neurology	23	46
HK	1	1995	Paediatric Neurology	10	10
HK	2	1979	Paediatric Neurology, professor	30	2
HK	3	2006	Paediatric Neurology trainee	2	3
P	1	1989	Paediatric Neurology	10	7.5
P	2	2002	Paediatrics	6	5
P	3	1989	Paediatric Neurology	4	20
P	4	1992	Paediatric Neurology	17	20
C	1	2000	Genetics attending	12	32
C	2	1996	Neurology attending	7	36
C	3	2006	Neurology attending	4	4
C	4	2010	Clinical Genetics fellow	2	4

C = Childrens' Hospital of Philadelphia; HK = the University of Hong Kong; Mu = Ludwig Maximilians Universität München; P = University of Pretoria; R = ErasmusMC Rotterdam

## Supplementary Document 14

### Intra-rater reliability for the IPMDS.

Intra-rater reliability ( $ICC_{\text{agreement}}$ ) per item of the IPMDS for two physicians scoring their own videotaped consultation. Items with an  $ICC_{\text{agreement}} \geq 0.7$  (good agreement) are marked in green;  $\leq 0.3$  (poor agreement) marked in blue. The items not suitable for video rating were removed.

ICC-agreement		
2.03	Alertness	NV
2.05	Dysarthria	0.769
2.06	Ptosis	NV
2.07	Strabismus	NV
2.08	Eye movements	0.000
2.09	Nystagmus	1.000
2.13	Hypokinesia	1.000
2.14	Abnormal movements	0.848
2.15	Ataxia	0.100
2.16	Tremor	NV
2.17	Reflexes	0.571
2.18	Hypertonia	0.897
2.19	Hypotonia	NV
2.20	Rigidity	0.935

NV = no variance

ICC-agreement		
3.01	Communicating	NV
3.02	Head control	NV
3.03	Rolling over	NV
3.04	Sitting up	-0.154
3.05	Sitting up	NV
3.06	Standing up	0.956
3.07	Standing	0.964
3.08	Walking	0.981
3.09	Running	0.968 x
3.10	Hopping	0.971 x
3.11	Reaching	0.833
3.12	Grasping	0.520
3.13	Rotating	0.938 x

X = only above 6 years; NV = no variance

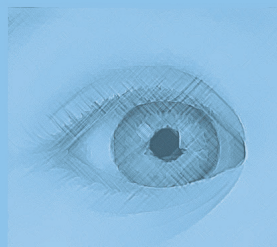
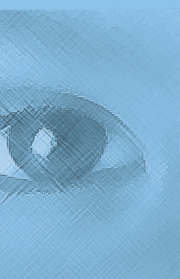
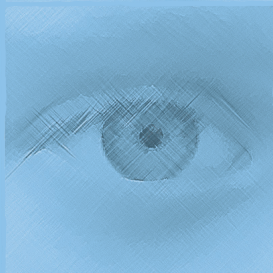
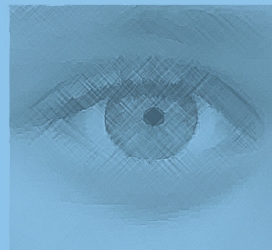
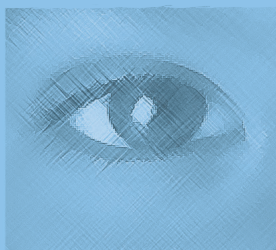
## Supplementary Document 15

### Test-retest reliability for the IPMDS.

Test-retest reliability ( $ICC_{\text{agreement}}$ ) per item of Domain 1 of the IPMDS for retest by the same rater by telephone and by a different rater (nurse specialist). Items with an  $ICC_{\text{agreement}} \geq 0.7$  (good agreement) are marked in green;  $\leq 0.3$  (poor agreement) marked in blue.

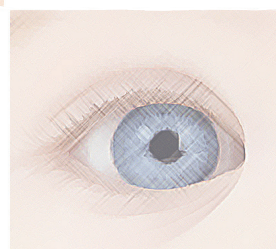
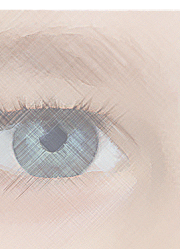
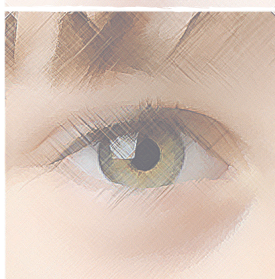
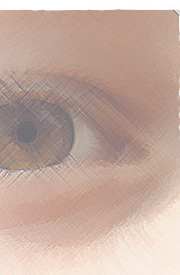
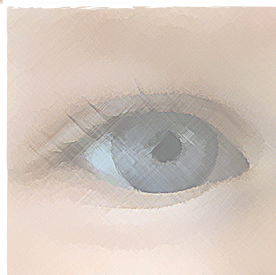
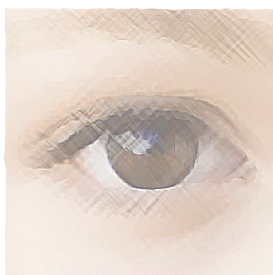
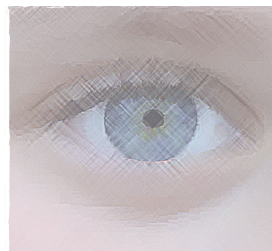
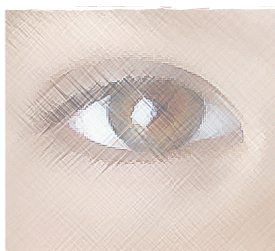
		Same rater ( $n = 4$ )	Different rater ( $n = 4$ )
1.01	Response to environment	( $n = 4$ )	0.571
1.02	Exercise capacity	1.000	0.451
1.03	Walking distance	1.000	1.000
1.04	Tiredness	1.000	0.348
1.05	Depressed	NV	0.000
1.06	Epilepsy	1.000	0.000
1.07	Headache	NV	0.727
1.08	Muscle pain	NV	0.923
1.09	Infections	NV	-0.465
1.10	Chewing	1.000	0.979
1.11	Swallowing	1.000	1.000
1.12	Hearing	NV	NV
1.13	Communicating	1.000	0.211
1.14	Vomiting	1.000	0.960
1.15	Gastroesophageal reflux	NV	0.800
1.16	Constipation	1.000	0.571
1.17	Diarrhoea	NV	NV
1.18	Cognitive functioning	1.000	0.000
1.19	Behavioural problems	NV	NV
1.20	Autistic features	1.000	0.800
1.21	Breathing subjective	1.000	1.000
1.22a	Continence day	1.000	NV
1.22b	Continence night	1.000	1.000
1.23a	Strabismus when tired	NV	
1.23b	Ptosis when tired	0.571	
1.23c	Dysarthria when tired	NV	

NV = no variance





# Discussion



# General discussion

*Partially in press by  
EMBO Molecular Medicine*

The aim of this thesis was to contribute to the knowledge on outcome measures in paediatric mitochondrial disease. This knowledge is essential to make rational decisions on which instruments to use in future clinical studies. This aim led to the following research goals: i) explore which symptoms and disabilities associated with mitochondrial disease should be measured in clinical trials from a patients' and doctors' perspective; ii) systematically select instruments from literature that could be used to cover these most important symptoms and disabilities; iii) validate these instruments in patients with mitochondrial disorders; and iv) develop new instruments for domains, which seem important but are not covered by instruments found during the systematic review.

Here, I will first elaborate on the methodology used to select the outcome measures and study populations for our validation studies. Subsequently, I will critically review the results of the validation studies and use these results to make recommendations on which instruments seem most promising. Finally, I will give suggestions to improve the quality of the selection process of outcome measures for future clinical trials in these rare disorders, as well as some other recommendations to improve the quality of such trials.

## Approach to the validation of outcome measures

In blockbuster disease clinical trials, in for example cardiovascular disease or cancer, time-to-event (relapse, stroke, myocardial infarction, death), histological results or radiological results are often used as end points.<sup>1-3</sup> Since at this time no markers for mitochondrial dysfunction are known that correlate with disease progression<sup>4</sup> (and Chapter 5, 10 and 11) and studies using time-to-event as their primary end point require the inclusion of large populations, the orphan disease field is faced with specific difficulties as it comes to the selection of outcome measures.<sup>5</sup> Since the number of patients to be included in a clinical trial (sample size) depends on the size of the effect relative to the precision of the measurement, either large treatment effects or sensitive and reliable outcome measures are needed when only a small amount of patients is eligible to participate in the study. Especially since therapies are expected to slow disease progression rather than reverse morbidity,<sup>6,7</sup> there is a huge challenge to identify instruments able to reliably quantify more subtle changes in for example strength, gait, and endurance.

Developing outcome measures for the paediatric population is challenging for many more reasons. Not only are children developing and growing (requiring age-specific tests and the use of age- and size matched reference values), their understanding and enthusiasm for and cooperation during the test may highly influence the test results. The feasibility of measuring disease progression in severely mentally disabled children,

whom are sometimes not even able to communicate effectively with their environment, let alone respond to an instruction, is limited. Since virtually all functional tests require the child to show the best (s)he can do, even children having milder mental disabilities might not show consistent results because of attention deficit and difficulties understanding the relevance of the tests. The measured results are therefore not only dependent on the stage of disease or grade of disability, but also on age, gender, intellectual functions, environment and personal factors.

## Selection of outcome measures

In part I of this thesis I aimed to select clinically relevant and robust outcome measures for future validation studies in children with mitochondrial disease.

Since this had never been done for children with mitochondrial disease, we used the experience from the neuromuscular field and cerebral palsy (CP) fields as an example. Both fields have a long history in measuring the effects of interventions on functional status. In the early 70's, the first study with steroids in boys with Duchenne muscular dystrophy (DMD) was performed.<sup>8</sup> At that time, only subjective and unstandardised activity assessment and strength measurements were available to evaluate the effect of treatment. Since this was felt to hamper the interpretation of the study results, the field quickly started with the development of standardised protocols including functional tests<sup>9-12</sup> and first optimisations in methodology.<sup>13,14</sup> Since then, many clinical studies investigating natural history and therapeutic interventions using different sets of outcome measures and populations have been performed. In 2011, the experience of eight centres, representing the longitudinal natural history data as well as experiences with the feasibility and reliability of various outcome measures in nearly 1,900 Duchenne patients over a 20-year time period, was presented at an international expert meeting.<sup>15</sup> These results were used to reach consensus for age-appropriate clinical outcome measures for clinical trials in boys with DMD.<sup>15</sup> To select the most valuable outcome measures in an iterative process, the neuromuscular community plies the following steps:<sup>16</sup> i) identification of the existing instruments; ii) choosing the optimal outcome measure(s) which is (are) both widely applicable and measures disease specific, clinically relevant (functional) capacities; iii) determine the minimal important difference; iv) determine the statistical properties in the disease population; v) confirm acceptance of the outcome measures by the regulatory authorities; and vi) listen to patients' and parents' experiences of the impact of the changes in functional abilities on their daily life.

In the CP field, patients and their caretakers were asked to rate the symptoms and limitations they would like to change from a much earlier stage.<sup>17,18</sup> Since the limitations caused by CP range widely and most interventions aim at targeting a specific symptom such as spasticity, the available instruments cover a broad pallet of body functions and

activities.<sup>19</sup> Several reviews have studied their psychometric properties<sup>20-24</sup> as well as the construct measured.<sup>25</sup> Although these reviews facilitate the selection of outcome measures targeted on a specific research question, consensus on which outcome measure to use for the most important domains still lacks.<sup>19,25</sup>

Our outcome measure identification project aimed at selecting psychometrically robust instruments measuring those aspects of the disease that are highly relevant to patients with mitochondrial disorders. The instruments were selected based the most burdensome symptoms according to patients and their parents and on the results of validation studies in DMD and CP, which were used as a model for mitochondrial myopathy and mitochondrial encephalopathy, respectively. One should however keep in mind that despite the similarities in functional disabilities, the underlying disease pathophysiology are completely different and this is naturally reflected in the clinical phenotype.

Similar to mitochondrial encephalopathy, the clinical picture as well as the underlying pathophysiology of CP is extremely heterogeneous,<sup>26</sup> including pyramidal syndromes, extrapyramidal syndromes, seizure disorders, and/or cognitive impairment.<sup>27</sup> Even symptoms rated as highly burdensome by mitochondrial disease patients, such as speech and language problems, and behavioural problems (Chapter 2), as well as oromotor dysfunction, hearing and vision problems, and attention deficits, are also seen in CP.<sup>27</sup> However, CP is by definition non-progressive,<sup>27</sup> while most mitochondrial encephalopathies have a progressive disease course.<sup>28,29</sup> The differences in pathophysiology are probably also reflected in e.g. in the aetiology of (fluctuations in) seizures<sup>30</sup> and cognitive capacities and in the patterns of daily activities performed by both groups.<sup>31</sup> Besides, specific syndromic associations, such as the brainstem and basal ganglia lesions in Leigh syndrome or the stroke-like episodes in MELAS syndrome, as well as the multi-system nature of mitochondrial disease are not reflected in the CP-model.

For DMD, the differences with mitochondrial myopathy are even clearer. Patients with DMD have a structural muscle problem, caused by (near-)absence of dystrophin, and thereby loss of mechanical stability during muscle contraction. This results in fragile muscle fibres that are prone to contraction-induced injury and progressive fibrosis.<sup>32</sup> As a consequence, boys with DMD suffer from progressive muscle weakness, which has a more or less predictable sequence and rate if untreated.<sup>33-38</sup> Outcome measures for DMD therefore mainly focused on muscle histology,<sup>39</sup> muscle strength,<sup>10,40</sup> and motor function,<sup>40,41</sup> whereas the limitations experienced by children with mitochondrial myopathy are mostly related to muscle endurance or repetitive movements. Still, the results of psychometric studies in DMD are of value since an instrument which is not able to reliably measure muscle power or functional abilities in children with DMD is unlikely to be a good instrument for patients with mitochondrial disease. Although it is possible that we removed promising instruments for mitochondrial disorders because of low

reliability scores in DMD in Chapter 3, our personal impression during the selection process was that none of the discarded instruments seemed more promising than the instruments selected for the mitochondrial toolbox.

## Study population

In part II of this thesis I aimed to determine inclusion criteria for our validation studies, based on the natural disease course and clinical heterogeneity in patients with nuclear encoded Complex I deficiency and in carriers of the m.3243A>G mutation. For our biomarker validation studies, we intended to include a large population with a similar genetic, biochemical and/or clinical background to minimise the influence of covariates. For the IPMDS and the accelerometer study we aimed to include a heterogeneous population to maximise generalizability of the feasibility and validity results.

The study on the natural history of nuclear encoded Complex I deficiency (Chapter 4) revealed that classification based on genetic defect, biochemical abnormalities, and clinical syndrome all led to small subgroups with still substantial clinical heterogeneity. Neither classification seemed superior in selecting homogeneous groups with a sufficient number of patients. In Chapter 4, we even found a substantial clinical heterogeneity in patients with Leigh syndrome caused by nuclear encoded Complex I deficiency, based on the localisation of the central nervous system abnormalities, the pattern of multi-system involvement and the reported disabilities. A recent study in patients with mutations in *SURF1*, an assembly gene for Complex IV, stated that *SURF1* patients have a homogeneous phenotype.<sup>42</sup> However, based on a careful evaluation of the data presented in this article and other reports,<sup>43</sup> we conclude that the clinical heterogeneity of *SURF1* patients equals to nuclear encoded Complex I deficiency. For this reason, and because of the low prevalence, we chose not to perform our surrogate marker validations in nuclear encoded Complex I deficiency or Leigh syndrome.

Recent m.3243A>G cohort studies by our (Chapter 5) and other groups<sup>44-46</sup> have illustrated the clinical heterogeneity which may be observed as a consequence of only one single nucleotide change in the mitochondrial DNA. We were able to recruit a total number of 82 carriers for this study within the first year and this number has increased to ~150 in 2014. Because of the large number of patients in which disease severity was assessed as part of the cohort study, we chose to validate our candidate surrogate measures in the cohort of m.3243A>G carriers. The availability of such a large genetically homogeneous cohort with known clinical phenotype outweighs the drawbacks of the (well-characterised) clinical heterogeneity.

## Validation Studies

In Chapter 3, we composed a mitochondrial toolbox of promising outcome measures based on the most burdensome complaints experienced by patients and their parents (Chapter 2) and the psychometric results in e.g. CP and DMD. This toolbox has been used as a starting point for the validation studies in part III of this thesis, which aimed to obtain experience with outcome measures in patients with mitochondrial disorders.

Several validation studies have been completed and included in this thesis. Although only very small groups with heterogenous phenotypes (children) or genetically homogeneous and clinically heterogenous (adults) populations were included and generalizability may be limited, the results of these validation studies have given an indication on which instruments are more promising compared to others. Here, I will elaborate further on the results of these validation studies.

### Functional outcome measures

Since tiredness and lack of energy were among the most burdensome symptoms according to patients and their parents, we sought for methods to quantify tiredness among the wide spectrum of children with mitochondrial disease. Since full capacity measured in a laboratory situation is highly influenced by personal factors such as motivation or anxiety and does not (always) reflect the disabilities the child experiences in daily life (e.g. sustained or repetitive movements), we aimed to measure tiredness in the patient's home situation using accelerometry (Chapter 6). We hypothesised that in children with mitochondrial disorders, the amount of physical activity reflects the exercise tolerance and the amount of energy the child is able to spend during the day. We expected a wide variability in activity during the day,<sup>47,48</sup> but hypothesised that the maximum peak-intensity and the total intensity of physical activity would be quite stable over the days. Seventeen children with mitochondrial disease and sixteen age- and gender-matched controls were included in this study. We found that patients had a lower peak intensity and that patients were resting more compared to their peers. These results seem to be generalizable, since children with widely variable symptoms and disabilities were included. In four patients and four healthy controls, the results of the retest in summer were not significantly different from the results in spring. When the technical issues we encountered during our validation study are addressed and the measurement protocol has been standardised to limit the effect of covariates, measuring daily physical activity using accelerometry may be of great importance in objectifying improvements in subjective observations of the patient such as feeling better and more energetic.

Exercise tests are obviously more closely related to the exercise capacity than daily physical activity. Since it is known that children – and especially those with disabilities



- rarely engage in maximal exercise,<sup>49</sup> and maximal and submaximal exercise capacity are closely related in healthy individuals,<sup>50-52</sup> we chose to evaluate a test for submaximal exercise capacity: the assisted 6-minute cycling test for arms and legs (A6MCT; Chapter 8). The validation study showed that the A6MCT seems to be feasible in children with OXPHOS dysfunction and a developmental age above approximately 4 years, without severe behavioural problems. This age-range is similar to the age at which the 6-minute walking test (6MWT) is used,<sup>53</sup> though the rate of completion of the 6MWT is known to rise dramatically after the age of 5 years.<sup>54</sup> The A6MCT was feasible in seven out of nine children, but none of the children included in our study seemed to need the assistance provided by the bicycle, as shown by the low number of abnormal results, in contrast to children with neuromuscular disease. Although the test was originally developed to replace the 6MWT in non-ambulatory children or children who are about to lose their ambulancy, the study design dictated that only ambulatory children could be included. Therefore, no solid conclusions about the feasibility and validity in non-ambulatory patients or patients with less muscle power can be drawn. Since only very little children with mitochondrial disease are both fully non-ambulatory and able to reliably complete a 6-minute task, confirming the reliability of the 6MWT in ambulatory children with mitochondrial disease who have mild to severe intellectual disabilities has higher priority.<sup>55</sup>

### General mitochondrial disease severity

Since some therapeutic effects are not known prior to the intervention, a broad screening instrument rating virtually all aspects of mitochondrial disease severity was developed: the International Paediatric Mitochondrial Disease Scale (IPMDS; Chapter 7). After a pilot reliability study involving eight patients and four raters, this scale was optimised and validated by 19 raters in 17 children with variable ages and phenotypes at five mitochondrial centres across the globe. IPMDS execution was feasible in all children, although both parents and physicians experienced difficulties in applying the IPMDS to toddlers, which was reflected in the inter-rater reliability of the second domain, but not in the first and third domain or in the number of items completed. The validation study showed that the IPMDS has good construct validity and a good reliability for most items. Excellent inter-rater reliability was found for the items within the functional assessment part and poor inter-rater reliability was found for 16 items (26%), mostly those included in the physical examination. More experience with the IPMDS across different age groups and mitochondrial phenotypes as well as knowledge about the minimal clinically important difference has to be gathered to facilitate the successful use of the IPMDS in future clinical studies.

## Surrogate markers

Since all above-mentioned functional measures are subject to patient bias (including e.g. voluntary effort and report bias), measuring disturbances of physiology may seem more objective and reliable. Biomarkers or surrogate measures are generally quick and easy to measure, objective, reliable, non-invasive, and possibly applicable to the full spectrum of patients with mitochondrial dysfunction.<sup>56</sup> Besides, biomarkers are not dependent of voluntary effort and can be measured reliably in severely affected patients of any age.

To date, none of the existing biomarkers for mitochondrial disease have been shown to correlate closely with a clinical outcome in any of the high-quality intervention studies.<sup>57</sup> In Chapter 10 and 11, we have evaluated the value of two potential new biomarkers for mitochondrial disease: Fibroblast Growth Factor 21 (FGF21) and growth and differentiation factor 15 (GDF15). In a large cohort of m.3243A>G carriers, we found only a moderate correlation between the concentration of both biomarkers and clinical disease severity, but neither of the two biomarkers was able to predict or monitor disease progression during follow-up. Because of the lack of correlation with disease progression, we could not yet recommend the use of FGF21 and GDF15 as an end point in clinical trials in m.3243A>G carriers. However, since m.3243A>G carriers only show minimal disease progression over the period of measurement and disease progression was measured using a clinical and partly subjective scale, it might still be of value to investigate the responsiveness of GDF15 and FGF21 during therapeutic interventions.

Two-dimensional speckle-tracking echocardiography (2DSTE; Chapter 9) was validated in our cohort of m.3243A>G carriers using data obtained from clinical care. Strain analysis was feasible in 90% of the images for global longitudinal strain (GLS). GLS was abnormal in 53 – 68% of cases (variation depending on reference values used). Drawing solid conclusions on the responsiveness of 2DSTE in m.3243A>G carriers was not yet possible because no structured clinical follow-up assessment was performed. The data for this retrospective pilot study were collected as part of routine care (using a standardised imaging protocol) and only 77% of the patients were suitable for the assessment of longitudinal myocardial strain using 2DSTE. Since these numbers are probably higher in highly standardised prospective studies, we recommend testing the feasibility and responsiveness of GLS in more detail in a prospective (intervention) study. Since abnormalities in myocardial deformation are frequently observed in patients with mitochondrial disease<sup>59-61</sup> and the 2DSTE method is a widely available and non-invasive method to quantify the performance of a non-voluntary muscle, global longitudinal strain is a promising end point for future clinical trials. To increase reliability of the results, the images should be analysed by two independent and experienced raters not aware of the condition and the treatment status of the carrier and efforts should be taken to standardise e.g. medication intake and resting conditions at the time of the

echocardiogram. Although the clinical relevance of subtle improvement in global longitudinal strain is not yet known, we hypothesised that an improvement in myocardial deformation reflects a decreased cardiomyopathy-associated morbidity and mortality.

## What outcome measure to use in future clinical trials?

The selection of appropriate outcome measures depends on many factors including the research goal, the type of the intervention, and the inclusion criteria (including age and mental capacities and functional abilities of in the intended patient group).<sup>25</sup> While early proof-of-principle studies would prefer to study a more narrowly defined population using highly detailed end points, larger trials require a larger population and thus more generally applicable outcome measures as to not render recruitment to a challenge.<sup>16</sup> Having seen over ~50 children and ~150 adults with mitochondrial disease in a structured program, either as part of clinical research or clinical care, we've begun to appreciate which tests might be useful and which instruments are certainly not. Here, I will elaborate further on which classes of instruments now look more promising and are of higher priority.

### Generic or disease/symptom specific instruments?

Concepts measured can be split roughly into generic (wellbeing, functioning) or disease-specific (symptom) aspects of the disease.<sup>62</sup> Since the spectrum of mitochondrial disorders is so heterogeneous, a non-categorical approach, addressing limitations in daily functioning seen as common manifestations of impaired health rather than specific to particular aetiologies, may suit best. We have applied this approach in a broader perspective using the ICF-CY to determine the domains to be included in our outcome measure inventory. Such a 'one-size-fits-all' approach increases the number of patients for which the instrument is relevant. However, in order to detect subtle, but meaningful changes in disease-specific symptoms by a generic instrument, large numbers of participants are required, especially in heterogeneous populations.<sup>16</sup> Two possible solutions for this dilemma include combining general instruments with disease-specific and symptom-specific instruments and the selection of a more homogeneous population for a clinical trial. The combination of generic and symptoms-specific outcome measures allows the inclusion of a broad clinical spectrum based on the relevance of the generic instrument, without losing the sensitivity of the organ-specific outcome measures for post-hoc analyses. The second approach may comprise the inclusion of a clinically homogeneous mitochondrial syndrome, such as Leber's Hereditary Optic Neuropathy. For these disorders, more targeted sets of outcome measures can be defined, although the prevalence of many of these disorder will make recruitment challenging and requires international collaborations (with its own challenges).

I suggest the use of a combination of generic instruments and a broad pallet of symptom-specific instruments in early phases of drug testing, to evaluate the effect of the drug both on in-depth organ function and on universal aspects such as functioning or quality of life. In subsequent efficacy studies in a more homogenous population, a smaller set of outcome measures may be applied, selected based on earlier trial results and adapted to the paediatric population.<sup>63</sup>

### Functional or surrogate measures?

Since all functional outcome measures are to some extent subjective instruments, either to the patient, his parents or to the investigator, one may feel a demand for more objective outcome measures that can be assessed using more reliable and unbiased assays or devices. These so-called surrogate end points are easy to measure variables that give an indirect indication of the toxicity or therapeutic benefit of a compound. Although surrogate end points have many advantages, such as a reduction of sample size and the possibility to indicate changes in biochemical processes that would be unethical to measure otherwise, drawbacks are numerous.

First and foremost, a change in a surrogate end point does not necessarily reflect a clinically meaningful effect of the treatment to the patient. In previous studies in progressive neuromuscular disorders, clinically relevant changes in motor function are not always equally reflected by changes in muscle power.<sup>64,65</sup> Moreover, improvements spasticity, joint contractures, gait dysfunction and bone displacements may not be reflected in functional abilities.<sup>66</sup> In mitochondrial disorders, examples include the lack of correlation between e.g. Complex I activity and the severity of the clinical phenotype (Chapter 4), the disappearance of lactate peaks on magnetic resonance spectroscopy during disease progression,<sup>4</sup> and the lack of meaningful correlation between changes in biomarkers FGF21 and GDF15 and disease progression (Chapter 10 and 11). The current view is that only functional outcome measures are good indicators of the aspects of function that reflect everyday activities.<sup>16</sup>

The second drawback of surrogate end points is that technical problems may affect the measurements. As shown by the accelerometer and 2DSTE study, instruments should be also tested in the intended population, since some seemingly easy tests may not be feasible in for example severely disabled children. Besides, fluctuations due to seasonal variation, physiological maturation and development, diurnal rhythm, hormonal cycle, food intake, exercise and local laboratory variations (measurement variability, storage, timing) should be mapped in great detail prior to using it in a clinical trial.

So, although surrogate outcome measures may serve as reliable and unbiased indications of proof of concept, their reflection of the final goal of the treatment, namely to improve the functional abilities or survival of the patient, should be established before

using it in efficacy studies.<sup>67</sup> Only sensitive and responsive biomarkers, closely related to a valid clinical end point, may suit as a surrogate end point. Moreover, changes in surrogate end points should always be supported by changes in clinically relevant parameters.

### Capacity or performance?

When choosing a functional instrument, it is important to distinguish between instruments measuring ((sub)maximal) capacity and those measuring the performance in daily life. The capacity is the ability to perform a certain task in a laboratory environment (e.g. on a level floor), while performance describes what an individual can do and does in his own environment on a daily basis (with both barriers and available aids and help).<sup>68</sup> Both constructs have their own advantages and disadvantages. For example, motor performance levels only partly reflect the motor capacities of the child,<sup>68</sup> and contextual factors, such as the physical and social environment and personal factors (anxiety, motivation) are likely to play a significant role. On the other hand, a laboratory situation in which a patient is encouraged to make a single maximum effort in a clinical environment by a relative stranger, is likely not to fully reflect the disabilities experienced in daily life.

Based on the results of our validation studies, including both submaximal exercise capacity tests and performance tests, I suggest including performance tests rather than capacity tests for patients with limited ability to follow instructions. For patients with more abilities, I suggest validating submaximal exercise tests, such as the 6MWT.

### Patient or clinician reported?

Both the Food and Drug Administration (FDA) and the World Health Organization (WHO) support the inclusion of patient-reported outcome measures to describe the patient's perspective on the effectiveness of the drug.<sup>62,69</sup> Health related quality of life (HRQOL), either generic or disease-specific,<sup>70</sup> is the most frequently used patient reported outcome measure. HRQOL is known to be only weakly to moderately correlated to physiological and functional parameters,<sup>71-73</sup> suggesting both that HRQOL instruments provide a unique perspective on the outcome of treatment but also that the universally applicable HRQOL questionnaires are influenced by many more factors than disease severity alone. Although measuring HRQOL seems like an art rather than science compared to measuring blood pressure, most HRQOL questionnaires show better reliability than diastolic blood pressure.<sup>74</sup> Nonetheless, since the paediatric field is highly dependent on proxy-report, and parents experience the child's quality of life different than the child does,<sup>36,38,75,76</sup> the validity of patient reported outcomes for this population seems limited.

Of course, the main question for the subjects included in a clinical trial is: 'In your opinion, does your treatment (or placebo) work? And how do you feel different?'. In my opinion, we should keep asking that question, especially in early phase drug testing, to be able to select outcome measures that test the changes reported by the patients. Patient reported outcome measures should only be included in a clinical trial after critical evaluation of clinical relevance of the content and reliability.

## Recommendations

Since different tools will provide unique and relevant information, it is expected that the inclusion of more instruments in the trial will be needed to gather sufficient information on the clinical condition of the patient. Naturally, the size of the set of outcome measures and the sequence of the tests should be adapted to the abilities and energy level of the patient. One should also realise that many tests are confronting to patients and their parents and therefore 'less is more'.

With the current knowledge, I would recommend the accelerometer as one of the main outcome measures in future clinical trials. This outcome measure seems most widely applicable since it measures a complaint relevant to a large proportion of children and does not require the child to follow specific instructions. Moreover, since the construct measured is closely related to the pathophysiological process that is reversed by the treatment, the end point is likely to respond to treatment effects. Before accelerometry can be used in clinical trials though, the hardware problems (*i.e.* battery of the devices) and the analyses need to be optimised. In parallel, patients and parents should be involved in defining the primary outcome and minimal important difference for this end point.

As secondary outcome measures, I propose (Box 1) the IPMDS, to assess general mitochondrial disease severity and screen for symptom-specific improvements, and 2DSTE, to objectively measure the function of a non-voluntary muscle. Since the Gross Motor Function Measure (GMFM) was feasible in all children included in the accelerometer study, I would suggest to monitor changes in gross motor abilities using the GMFM. The Paediatric Evaluation of Disability Inventory (PEDI) was feasible in all children seen both in the accelerometer study and in the IPMDS study, but since the PEDI has a significant ceiling effect in children with more abilities, I would only suggest this test only in more severely affected children. As a formal endurance test, I suggest the 6MWT, since most of the children who are able to follow instructions will be ambulatory. The functional abilities tests used as part of the A6MCT validation could be useful, but should be highly standardised and tailored to the population which will be included in the clinical trial. For the children in the A6MCT study (with a high level of abilities), we sometimes felt as if the results were dependent on the reaction time of the rater, instead of the abilities of the child.

**Box 1: Recommended tests for future clinical trials in paediatric mitochondrial disease\***

Test	Domain	Prerequisites
Accelerometer	Physical activity	Solve hardware problems
		Define more detailed primary outcome
		Define minimal clinically important difference
IPMDS	General mitochondrial disease severity	Define minimal clinically important difference
		Improve inter-rater reliability for Domain 2
2DSTE (global longitudinal strain)	Myocardial deformation	Assess feasibility in a prospective, highly-standardised study
		Determined by 2 independent raters not aware of the clinical condition of the subject
		Validate in children
		Standardise medication intake and resting conditions
		Define minimal clinically important difference
GMFM	Gross motor function	Test inter- and intrarater reliability, especially in children not able to follow instructions
		Define minimal clinically important difference
PEDI	Functional capabilities and performance	Only for severely affected children
		Determine test-retest and inter-rater reliability
6MWT	Submaximal exercise capacity	Determine test-retest reliability and learning effect
		Shorten if possible
		Define minimal clinically important difference
Timed function tests	(Speed of) functional abilities	Tailor to targeted population
		Increase level of standardisation
		Define minimal clinically important difference

Tests we cannot recommend (yet) include: fibroblast growth factor 21 (FGF21); growth and differentiation factor 15 (GDF15); Movement Disorder Childhood Rating Scale, Modified Tardieu test for spasticity, Motor Function Measure (MFM). 2DSTE = 2-dimensional speckle tracking echo-cardiography; 6MWT = 6 minute walking test; GMFM = Gross Motor Function Measure; IPMDS = International Paediatric Mitochondrial Disease Scale; PEDI = Pediatric Evaluation of Disability Inventory

\* Based on studies performed in this thesis

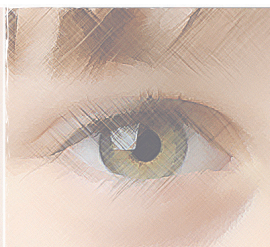
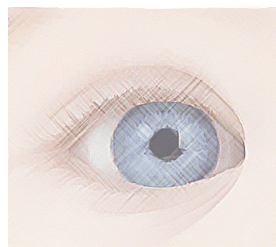
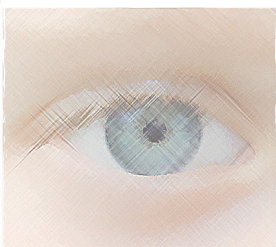
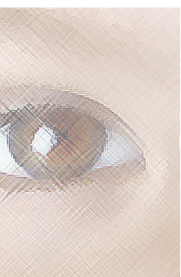
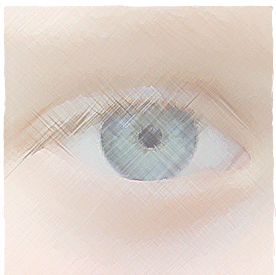
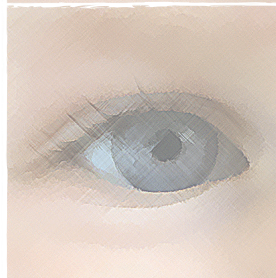
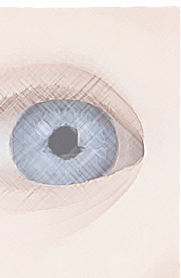
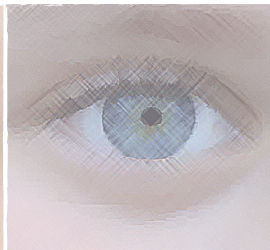
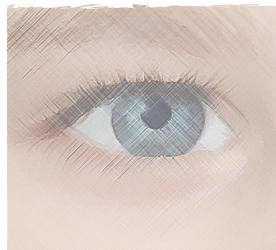
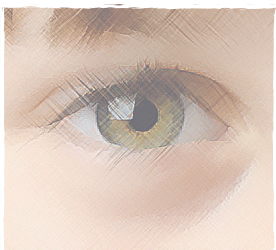
Several other tests included in our studies had lower suitability to paediatric mitochondrial disease compared to the above-mentioned tests. The Movement Disorder Childhood Rating Scale<sup>77,78</sup> was tested during the pilot study of the accelerometer. In our opinion, the test was not suitable for mitochondrial disorders, since the score on many items (e.g. handwriting, walking) was mostly affected by symptoms other than the movement disorder. Moreover, it is known that the inter-rater reliability of assessing

movement disorders is rather low.<sup>79</sup> The Modified Tardieu test for spasticity,<sup>80</sup> used as part of the accelerometer study, had good face validity but seemed to be quite situation- and rater-dependent. Especially since an improvement in spasticity is not always reflected in the functional abilities,<sup>66</sup> I don't recommend this test as a clinical end point for future clinical trials. The Motor Function Measure (MFM) was included in the A6MCT study. For this test, some of the items seemed highly relevant, but since many of the items were normal in most children, we don't recommend this test for children with mitochondrial myopathy. Some of the clinically relevant items of the MFM were included in the functional domain of the IPMDS.

A selection of the above mentioned tests, tailored to the population which is included in the study, may be complemented with surrogate markers and functional outcome measures (where indicated adapted for the paediatric population) selected from patient reports in earlier phase clinical trials. Since all of the above mentioned tests are, to a more or lesser extent, burdensome to patients and subject to either patient, parent or rater bias, I encourage the continuation of biomarker studies, to detect biomarkers correlating with clinically relevant organ functions and abilities.







Future prospects

Substantial progress has been made since this project started in 2010. Not only do we have more feeling for which outcome measures seem most robust, also several other initiatives focussing on the selection of outcome measures in mitochondrial disease have arisen, including the Common Data Elements for mitochondrial disease project from the National Institute of Health and a multidisciplinary workshop on biomarkers and outcome measures in Hinxton, United Kingdom. Beside the selection of outcome measures, many other challenges remain in designing a scientifically robust clinical trial protocol, including the selection of inclusion criteria and the choice of a primary and secondary outcome measures.

## Future Studies On Outcome Measures In Paediatric Mitochondrial Disease

Ideally, the selection of a test for use in a clinical trial is based on the experimental data on the test characteristics (e.g. feasibility, reliability, responsiveness and minimal important change) in the considered study population. As part of the 'Towards the Harmonisation of Outcome Measures for Children with Mitochondrial Disease'-project, we've designed the SO-MITO (Selecting Outcome measures for Mitochondrial disease) study (MREC nr NL.47373.091.13).

### The SO-MITO study

The SO-MITO study measures the clinical relevance, feasibility, reliability, validity, responsiveness, and minimum important change of a set of promising instruments from our toolbox in children with mitochondrial disease. Inter- and intrarater reliability are tested by assessing the patient with two blinded investigators and by making a video of the patient to be rated again by one of the investigators later. Test-retest reliability is assessed by asking the patient to perform the same tests within a limited time frame, two weeks later (if the patient is not feeling well or the disease seems not stable, this part of the protocol is postponed beyond the last visit). Validity is assessed by correlating tests measuring the same concepts, or by correlating with the patients' or rater's opinion about the concept measured (e.g. a VAS scale for general disease severity for patients or for motor abilities or muscle power for raters). Feasibility, clinical relevance and face validity are assessed during these study procedures. Responsiveness is tested by one of the raters at the time the intended clinical trial would finish. Since responsiveness requires the disease to change and there's no effective treatment available yet, responsiveness relies on the fluctuating disease course (both in a positive and negative direction) seen in many patients. Another suggestion would be to combine this study with an intervention study, e.g. aerobic training to prevent secondary disuse.<sup>61</sup> However, since most children are already under such a treatment and the burden of participation will be higher, this may not be desirable. Since knowing the

minimum important difference is essential before designing future clinical trials (e.g. for choosing the most responsive outcome measure and to determine the sample size), no effort should be spared to determine this parameter. The minimal clinically important difference can be assessed by for example assessing which changes patients and parents experienced as relevant compared to the previous assessment in a prospective natural history study or by statistically-based approaches, although the latter do not take into account the patients' perspective.<sup>34</sup> To tailor the SO-MITO study protocol to the to-design clinical trial, the protocol should be adapted to include the proposed study population and having the same length of time that will be used in the trial in order to assess the spectrum of possible changes in time.<sup>16</sup> Although the burden of this study to patients, investigators and financial resources seems limited, the burden associated with the execution such as study before the start of a clinical trial (which is preferably in the same patient population) may still be too burdensome to patients and their families.

To gather similar data in a more efficient way, many aspects of the above mentioned protocol may be integrated in a early phase clinical trial. Examples of protocol adaptations include: i) repeated baseline measurements to assess test-retest reliability and learning effects; ii) blinded assessment by two assessors and videotaping for rater-dependent tests in this kind of tests; iii) asking patients about their experience with the individual tests within the study protocol; and iv) asking patients to rate relevant changes in specific disease aspects or functional abilities to determine the minimum important difference. These adaptations will provide highly relevant information for future studies, but again may be an extra burden to patients already in a strenuous research protocol.

Another option is the use of outcome measures as part of clinical care. Although standardisation will be suboptimal, the raters will have a good feeling about which tests are feasible, reliable, valid and responsive. This may also form the basis on which outcome measures are selected for more intensive protocols such as the above mentioned SO-MITO study. In Nijmegen, we've been running the so-called 'MitoRoute' since 2014. In this intensive screening programme, about 40 different tests are used. We are currently analysing the results of the first 50 patients.

### Optimisation of outcome measures

It is likely that some tests seem ready to be in- or excluded in future studies, while others will have to be optimised to be feasible or more widely applicable to the intended population. Here, I will give some suggestions for adaptations of available tests, for newly to develop tests and of new approaches to define end points in intervention studies. Obviously, these tests will have to be validated and if indicated undergo a second optimisation round before they can be used to determine treatment effects.

The 6-minute chewing test is a new instrument under development by our speech and language therapists. Since it has been developed based on the observation that many children with mitochondrial disease were not able to sustain chewing for the full duration of a meal, this is a good example of a clinically relevant outcome measure. The duration of this test is again 6 minutes. The first results of this test show that children only deviate from normal within the last minutes. This is not the case for the 6MWT and the A6MCT, which were completed at a stable pace (Chapter 7). Since it is known that children with attention problems have difficulties completing the full 6MWT,<sup>53</sup> I suggest to test the validity of a 1-minute or 3-minute version of the A6MCT and 6MWT.<sup>82</sup> Besides, a training session before the start of the study should be considered, since a test-retest study of the 6MWT in patients with Down syndrome showed a substantial learning effect.<sup>55</sup> Moreover, one should take into account that, since children are growing and developing, motor abilities and muscle power are expected to be higher during follow-up. On the other hand, some child may fully lose the ability to perform the test during disease progression. Therefore, the result of these tests should be expressed as a z-score (normalised for height or age) or as velocity, instead of a distance or a time.

Developing an entirely new rating scale for motor function abilities in children for mitochondrial disorders, as has been done multiple times for DMD,<sup>83,84</sup> is complicated by the phenotypic heterogeneity of mitochondrial disorders. Separate scales for children with mitochondrial myopathy and mitochondrial encephalopathy may be suggested, although the need for a new functional scale for mitochondrial encephalopathy is low since the GMFM seems a solid test for this patient group. Except for the Childhood Myositis Assessment Scale,<sup>85</sup> I am not aware of any functional childhood scale testing muscle endurance. Although the Childhood Myositis Assessment Scale showed good psychometric properties<sup>85,86</sup> and measures an important aspect of mitochondrial myopathy, there is an important ceiling effect and children need to be cooperative and motivated to complete the test. Other suggestions for newly to develop scales include a paediatric mitochondrial disease quality of life scale, analogous to the adult scale,<sup>87</sup> where possible with limited proxy-report.

Other possible adaptations include goal attainment scaling, an individualised outcome measures to evaluate the achievement of pre-specified domain and goals (e.g. ‘I would like to eat without the assistance of my caretakers. I rate my improvement a 3 when I am able to eat fully independently, a 2 when I need assistance with soup, a 1 if I’m able to complete at least 50% of my meal without assistance and a 0 if I don’t improve at all.’) or combinations of instruments or summed items which can be used as a single end point. Judging treatment effects based on pre-defined changes in a set of personalised outcome measures facilitates the inclusion of larger, more heterogeneous populations into the study, without losing the ability to detect small but relevant treatment effects. These personalised adaptations may be extremely relevant in these complex multi-system disorders.

## Selection of outcome measures

After completion of the above mentioned studies, a set of outcome measures can be selected, based on the psychometric properties, relevance and burden to the patient, and feasibility in multi-centre studies (Box 2).<sup>16,69,88,89</sup>

### Box 2: The ideal outcome measure...

- ...reflects a clinically relevant outcome or change in disease severity
- ...has been validated for the specific study protocol (*i.e.* study population, time to second measurement)
- ...has acceptable test-retest reliability
- ...is sensitive to change
- ...has an acceptable patient burden
- ...serves a broad age group
- ...assesses a large range of abilities
- ...is relevant to the largest possible phenotypic spectrum
- ...has little cultural bias
- ...has little (or a known and stable) practice effect
- ...is not highly dependent on voluntary effort
- ...has natural history data available for intended population
- ...is easy to execute with a limited amount of training

Preferably, the choice for a specific (set of) outcome measure(s) is not made by a single doctor or research team, but by a team of key opinion leaders consisting of doctors, methodologists, regulators, industry representatives and patient - and parent representatives from various institutions, using a clear pre-defined protocol based on which decisions will be made.<sup>90</sup>

Since the low prevalence of mitochondrial disorder will probably require multi-centre, multi-national trials, involving international experts in an early stage is of main importance. When executing such an international trial, differences between sites should be minimised despite *e.g.* the large number of evaluators and language differences.<sup>91</sup> Therefore, special care should be taken to guarantee training for all raters and to test and diminish inter-rater reliability. If patients are too scattered around the world, one might think of relocating patients and their families to the study site or of a combination of centralised pre- and post treatment evaluations in combination with assessments which can be performed using telemedicine or devices to quantify functional abilities.<sup>92</sup>

International agreement on which outcome measures to use for specific populations is also indicated for the collaboration in natural history studies. More and more patients are diagnosed with genetic defects associated with mitochondrial disease. As illustrated in our study in nuclear encoded Complex I deficiency, there is a wide heterogeneity between patients with the same gene defects. The standardised evaluation of

these patients in natural history studies using clinically relevant outcome measures will provide both information for interpreting clinical trials as for communicating the prognosis in relevant terms to patients and families.

## Clinical Trial Design In Paediatric Mitochondrial Disease

The selection of outcome measures is just one of the aspects of designing a clinical trial. As for the selection of outcome measures, the selection of a study population and study design is highly dependent on the research question. For this discussion, I assume that a non-specific improvement in disease severity, energy levels and motor functioning is expected after application of the drug.

### Study population

To draw correct conclusions about the efficacy of treatment with a low number of patients (small sample size), the variance between subjects should be as low as possible. In other words, we would like to include a homogeneous cohort of patients, with similar genotype and phenotype, similar disease stage, similar organs involved and similar expected effect of treatment, without the use of other medication, liver- or kidney failure (one of the recommendations in Chapter 1). Phenotypically homogeneous mitochondrial syndromes with an acceptable prevalence (e.g. Leber's Hereditary Optic Neuropathy) are rare and none of these syndromes is paediatric. Moreover, because of the previously described challenge of selecting end points for children who are not able to communicate effectively or to follow instructions, I suggest that efficacy of drugs should be studied primarily in adults with mitochondrial disease. Since children are not small adults, the administration of drugs to children requires separate studies on formulation, pharmacokinetics and therapeutic effects. Although the efficacy of the drug is already proven when these paediatric studies are conducted, the quality of the study should still be sufficient to prove the drug's pharmacokinetics, safety and efficacy. To improve the quality of the efficacy study, I will make some recommendations regarding the study population.

Some rare, more or less homogeneous paediatric mitochondrial syndromes exist, such as Barth syndrome, a mitochondrial membrane disorder caused by a mutation in the TAZ gene. Barth syndrome seems to have a reasonably homogeneous phenotype with skeletal muscle weakness, cardiomyopathy, neutropenia, and growth delay.<sup>93-97</sup> Since Barth syndrome is a mitochondrial membrane disorder, the OXPHOS dysfunction may be classified as secondary to the membrane disruption. Kearns Sayre syndrome is another more or less homogeneous mitochondrial syndrome characterised by ophthalmoparesis, pigmentary retinopathy, deafness, muscle weakness, ataxia, and cardiac conduction block.<sup>98</sup> The progression of the disease seems to be fairly stable.<sup>99</sup> The fact



that this syndrome affects mostly mentally-competent children and adolescents facilitates the feasibility and the ethical acceptability of the study, although a separate study will have to be performed in neonates, infants and young children.

Another option is to include all children within a certain age group with mitochondrial disorders, regardless of their genetic or biochemical diagnosis, but with very strict clinical inclusion criteria. Examples of these would be for example: a stable disease course in the past year; able to follow instructions such as hopping on one leg or rotating a pen in the hand; a 6MWT between the 5th and 10th percentile for age or height; and/or a GMFM score between 15 and 25%.

Although measuring reliably in children is challenging, treatment effects may be expected to be even larger compared to adults because of higher neuroplasticity.<sup>27</sup> Especially in children, effects and side effects should be monitored closely after approval of the drug, to select patient groups in which the balance between effect and side effects is negative, not only to reduce the pressure on the financial resources<sup>100</sup> but also since long term effects of drugs are mostly poorly characterised.

## Study design

Since the classical randomised controlled clinical trials generally require large numbers of patients, especially in heterogeneous populations, several alternative trial designs are available.<sup>101-103</sup>

The use of a cross-over design, stepped-wedge, or a *n*-of-1 trial minimises variance and confounding and guarantees exposure to active treatment, which facilitates recruitment of patients. In all three designs, the patient undergoes both active treatment and placebo treatment and is used as its own control. Whereas cross-over and stepped wedge studies aim to determine the efficacy of a drug, an *n*-of-1 study aims to determine the effect of a drug within a single subject. The results of multiple *n*-of-1 studies can be used in a meta-analysis.<sup>103</sup> Cross-over, stepped wedge, and *n*-of-1 designs can only be applied if the disease status is expected to be similar between the start of the first and the second treatment phase, they can only be applied in clinically stable paediatric mitochondrial disorders. Moreover, if the duration of the intervention and the wash-out period are short and no long-term effects of treatment are expected (*i.e.* minimal carry-over effect), the cross-over design and *n*-of-1 trial designs may be of great value when confirming drug efficacy in mitochondrial disease. Where feasible, larger and phenotypically homogeneous cohorts can be analysed in cross-over studies. Choosing the methodologically less solid stepped wedge design is only acceptable if substantial long-term effects of treatment (carry-over) are expected. In rare mitochondrial syndromes or those with a heterogeneous phenotype, *n*-of-1 trials are most suitable. In both cases, patients with an obviously oscillating disease course should be excluded.

When choosing any of these designs, several other methodological issues should be covered. First of all, the sequence of the placebo and the active treatment phases should be randomised to correct for confounders such as seasonal influences and (unexpected) learning effects. In case of group comparisons, such as in a cross-over study, randomisation should be stratified based on predefined criteria, in order to create comparable groups. Secondly, since all of the above mentioned study designs are subject to report bias, blinding of both investigators and patients is of great importance, especially in case of subjective end points. Thirdly, the end points used in these trials should be harmonized over individuals and/or groups to be able to draw conclusions about the overall efficacy of the drug. Finally, since the outcome measure is applied multiple times within the same subject, the instruments should be tested for learning processes, which is expected to be the case for *e.g.* the A6MCT and the 6MWT.

Another adaptive trial design is the Bayesian trial design, which provides more clinically intuitive results. In contrast to the confidence intervals and  $p$ -values provided by the frequentist designs, this design provides probabilities that a predefined change occurs (*e.g.* the chance of a 10% higher GMFM after 4 weeks of active treatment compared to placebo is 23%). Although this design might still be unfamiliar to regulatory authorities and collaborators, its flexible design and required smaller samples sizes provides promising design for phase 3 trials.

## Conclusion

The field of clinical mitochondrial medicine has been gaining momentum in the past decade. We have accomplished that it is now widely accepted that patients shouldn't be bothered by low-quality studies. The responsibility of the international expert centres is now to set guidelines for high quality studies and to facilitate improvements in trial design, *e.g.* by harmonizing outcome measures. Naturally, the decision on which end points to use for natural history studies and clinical trials should be guided by the research question and taken by an expert panel and based on data obtained from experiments such as those in this thesis.

For drug trials in the near future, I would suggest to determine the efficacy of the drug in adults with mitochondrial disease first, followed by a high-quality paediatric study. To determine the efficacy of a drug in children, I would suggest either a randomised, placebo-controlled, double-blinded Bayesian design in all children with mitochondrial disease or a randomised, placebo-controlled, double-blinded cross-over study in a more homogeneous and clinically stable cohort of children. The efficacy can be determined either by the minimal important difference in a primary outcome measure or by pre-defined changes in a personalised set of end points. Both approaches require more experience with outcome measures in children with mitochondrial disorders, which is still an important topic of our ongoing research.

## References

1. Etzioni R, Gulati R, Lin DW. Measures of survival benefit in cancer drug development and their limitations. *Urologic oncology* 2015;33:122-7.
2. Fernandez-Martos C, Guerrero A, Minsky B. Clinically relevant study end points in rectal cancer. Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer 2012;196:3-19.
3. Preiss D, Sattar N, McMurray JJ. A systematic review of event rates in clinical trials in diabetes mellitus: the importance of quantifying baseline cardiovascular disease history and proteinuria and implications for clinical trial design. *American heart journal* 2011;161:210-9 e1.
4. Lee SK, Kim J, Kim HD, Lee JS, Lee YM. Initial experiences with proton MR spectroscopy in treatment monitoring of mitochondrial encephalopathy. *Yonsei Med J* 2010;51:672-5.
5. Augustine EF, Adams HR, Mink JW. Clinical trials in rare disease: challenges and opportunities. *Journal of child neurology* 2013;28:1142-50.
6. Kang PB. Beyond the gowers sign: measuring outcomes in Duchenne muscular dystrophy. *Muscle Nerve* 2013;48:315-7.
7. Friedman LS, Farmer JM, Perlman S, et al. Measuring the rate of progression in Friedreich ataxia: implications for clinical trial design. *Movement disorders : official journal of the Movement Disorder Society* 2010;25:426-32.
8. Drachman DB, Toyka KV, Myer E. Prednisone in Duchenne muscular dystrophy. *Lancet* 1974;2:1409-12.
9. Cornelio F, Dworzak F, Morandi L, Fedrizzi E, Balestrini MR, Gondoni L. Functional evaluation of Duchenne muscular dystrophy: proposal for a protocol. *Ital J Neurol Sci* 1982;3:323-30.
10. Scott OM, Hyde SA, Goddard C, Dubowitz V. Quantitation of muscle function in children: a prospective study in Duchenne muscular dystrophy. *Muscle & nerve* 1982;5:291-301.
11. Griggs RC, Moxley RT, 3rd, Mendell JR, et al. Prednisone in Duchenne dystrophy. A randomised, controlled trial defining the time course and dose response. *Clinical Investigation of Duchenne Dystrophy Group*
12. Clinical investigation of Duchenne muscular dystrophy. Interesting results in a trial of prednisone. *Arch Neurol* 1991;48:383-8.
13. Mendell JR, Province MA, Moxley RT, 3rd, et al. Clinical investigation of Duchenne muscular dystrophy. A methodology for therapeutic trials based on natural history controls. *Arch Neurol* 1987;44:808-11.
14. Florence JM, Pandya S, King WM, et al. Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. *Phys Ther* 1984;64:41-5.
15. Brooke MH, Fenichel GM, Griggs RC, et al. Clinical investigation in Duchenne dystrophy: 2. Determination of the 'power' of therapeutic trials based on the natural history. *Muscle Nerve* 1983;6:91-103.
16. Bushby K, Connor E. Clinical outcome measures for trials in Duchenne muscular dystrophy: report from International Working Group meetings. *Clinical investigation* 2011;1:1217-35.
17. Mercuri E, Mazzone E. Choosing the right clinical outcome measure: from the patient to the statistician and back. *Neuromuscul Disord* 2011;21:16-9.
18. Vargus-Adams JN, Martin LK. Domains of importance for parents, medical professionals and youth with cerebral palsy considering treatment outcomes. *Child: care, health and development* 2011;37:276-81.

19. Vargus-Adams JN, Martin LK. Measuring what matters in cerebral palsy: a breadth of important domains and outcome measures. *Arch Phys Med Rehabil* 2009;90:2089-95.
20. Vargus-Adams J. Understanding function and other outcomes in cerebral palsy. *Physical medicine and rehabilitation clinics of North America* 2009;20:567-75.
21. Debuse D, Brace H. Outcome measures of activity for children with cerebral palsy: a systematic review. *Pediatr Phys Ther* 2011;23:221-31.
22. Mitchell LE, Ziviani J, Oftedal S, Boyd RN. A systematic review of the clinimetric properties of measures of habitual physical activity in primary school aged children with cerebral palsy. *Res Dev Disabil* 2013;34:2419-32.
23. James S, Ziviani J, Boyd R. A systematic review of activities of daily living measures for children and adolescents with cerebral palsy. *Developmental medicine and child neurology* 2014;56:233-44.
24. Clanchy KM, Tweedy SM, Boyd R. Measurement of habitual physical activity performance in adolescents with cerebral palsy: a systematic review. *Developmental medicine and child neurology* 2011;53:499-505.
25. Capio CM, Sit CH, Abernethy B, Rotor ER. Physical activity measurement instruments for children with cerebral palsy: a systematic review. *Developmental medicine and child neurology* 2010;52:908-16.
26. Schiariti V, Klassen AF, Cieza A, et al. Comparing contents of outcome measures in cerebral palsy using the international classification of functioning (ICF-CY): a systematic review. *Eur J Paediatr Neurol* 2014;18:1-12.
27. MacLennan AH, Thompson SC, Gecz J. Cerebral palsy: causes, pathways, and the role of genetic variants. *American journal of obstetrics and gynecology* 2015.
28. Aisen ML, Kerkovich D, Mast J, et al. Cerebral palsy: clinical care and neurological rehabilitation. *Lancet Neurol* 2011;10:844-52.
29. Skladal D, Sudmeier C, Konstantopoulou V, et al. The clinical spectrum of mitochondrial disease in 75 pediatric patients. *Clin Pediatr (Phila)* 2003;42:703-10.
30. Scaglia F, Towbin JA, Craigen WJ, et al. Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics* 2004;114:925-31.
31. Rahman S. Mitochondrial disease and epilepsy. *Developmental medicine and child neurology* 2012;54:397-406.
32. Martens AM, Gorter H, Wassink RG, Rietman H. Physical activity of children with a mitochondrial disease compared to children who are healthy. *Pediatr Phys Ther* 2014;26:19-26.
33. Deconinck N, Dan B. Pathophysiology of duchenne muscular dystrophy: current hypotheses. *Pediatr Neurol* 2007;36:1-7.
34. Vuillerot C, Girardot F, Payan C, et al. Monitoring changes and predicting loss of ambulation in Duchenne muscular dystrophy with the Motor Function Measure. *Developmental medicine and child neurology* 2010;52:60-5.
35. Ciafaloni E, Fox DJ, Pandya S, et al. Delayed diagnosis in duchenne muscular dystrophy: data from the Muscular Dystrophy Surveillance, Tracking, and Research Network (MD STARnet). *The Journal of pediatrics* 2009;155:380-5.
36. Bartels B, Pangalila RF, Bergen MP, Cobben NA, Stam HJ, Roebroek ME. Upper limb function in adults with Duchenne muscular dystrophy. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine* 2011;43:770-5.

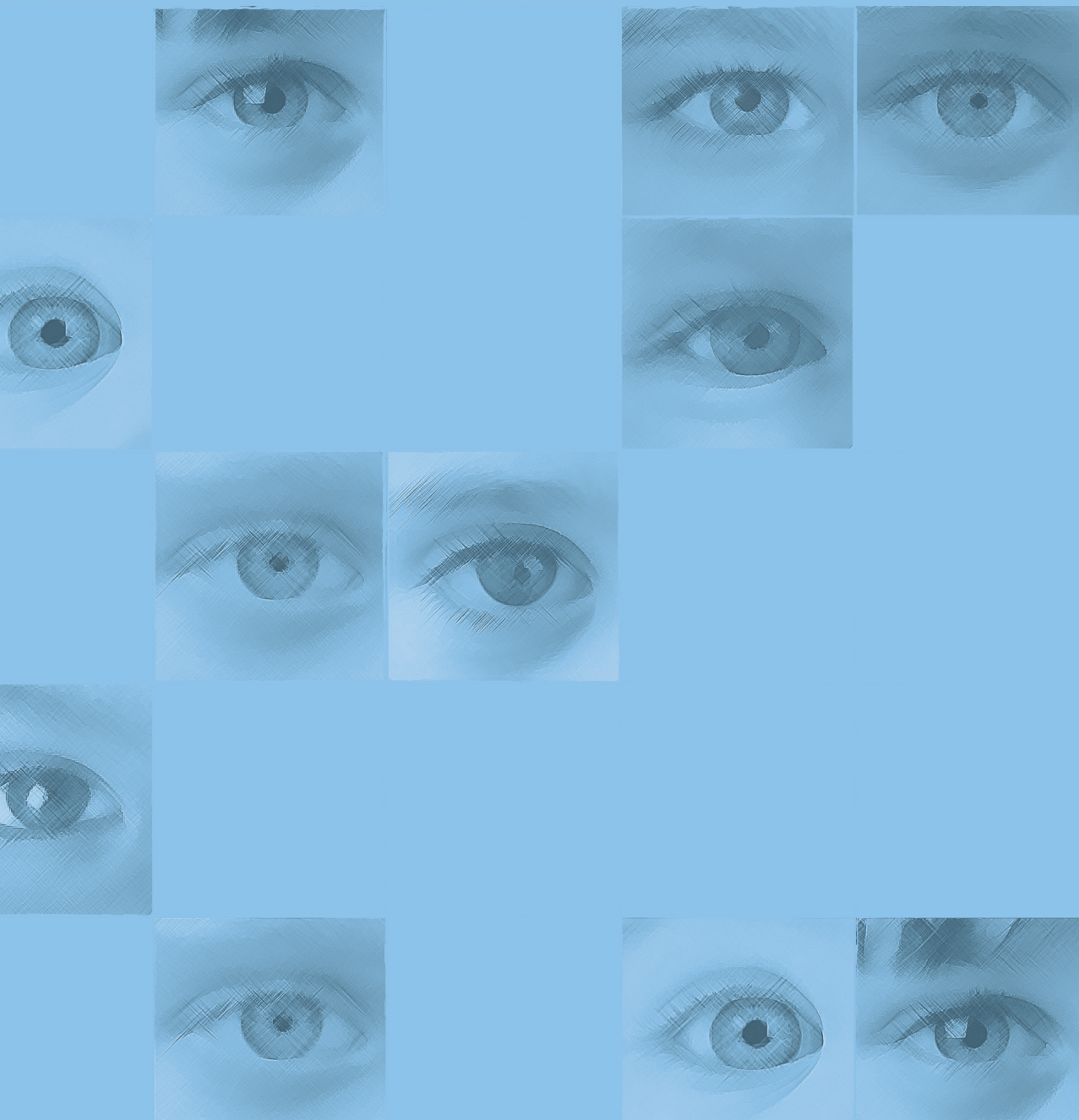
37. Bach JR, Martinez D. Duchenne muscular dystrophy: continuous noninvasive ventilatory support prolongs survival. *Respiratory care* 2011;56:744-50.
38. Kohler M, Clarenbach CF, Bahler C, Brack T, Russi EW, Bloch KE. Disability and survival in Duchenne muscular dystrophy. *Journal of neurology, neurosurgery, and psychiatry* 2009;80:320-5.
39. Pane M, Mazzone ES, Sivo S, et al. Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. *PloS one* 2014;9:e108205.
40. Jansen M, van Alfen N, Nijhuis van der Sanden MW, van Dijk JP, Pillen S, de Groot IJ. Quantitative muscle ultrasound is a promising longitudinal follow-up tool in Duchenne muscular dystrophy. *Neuromuscul Disord* 2012;22:306-17.
41. Parreira SL, Resende MB, Della Corte Peduto M, Marie SK, Carvalho MS, Reed UC. Quantification of muscle strength and motor ability in patients with Duchenne muscular dystrophy on steroid therapy. *Arq Neuropsiquiatr* 2007;65:245-50.
42. Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord* 2005;15:463-70.
43. Wedatilake Y, Brown RM, McFarland R, et al. SURF1 deficiency: a multi-centre natural history study. *Orphanet journal of rare diseases* 2013;8:96.
44. Aulbert W, Weigt-Usinger K, Thiels C, et al. Long survival in Leigh syndrome: new cases and review of literature. *Neuropediatrics* 2014;45:346-53.
45. Nesbitt V, Pitceathly RDS, Turnbull DM, et al. The UK MRC Mitochondrial Disease Patient Cohort Study: clinical phenotypes associated with the m.3243A > G mutation-implications for diagnosis and management. *J Neurol Neurosur Ps* 2013;84:936-8.
46. Kaufmann P, Engelstad K, Wei Y, et al. Natural history of MELAS associated with mitochondrial DNA m.3243A>G genotype. *Neurology* 2011;77:1965-71.
47. Ma Y, Fang F, Cao Y, et al. Clinical features of mitochondrial DNA m.3243A>G mutation in 47 Chinese families. *J Neurol Sci* 2010;291:17-21.
48. Jeannot PY, Aminian K, Bloetzer C, Najafi B, Paraschiv-Ionescu A. Continuous monitoring and quantification of multiple parameters of daily physical activity in ambulatory Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2011;15:40-7.
49. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. *Prev Med* 2007;44:143-7.
50. Cooper DM. Rethinking exercise testing in children: a challenge. *Am J Respir Crit Care Med* 1995;152:1154-7.
51. Lesser DJ, Fleming MM, Maher CA, Kim SB, Woo MS, Keens TG. Does the 6-min walk test correlate with the exercise stress test in children? *Pediatr Pulmonol* 2010;45:135-40.
52. Li AM, Yin J, Yu CC, et al. The six-minute walk test in healthy children: reliability and validity. *Eur Respir J* 2005;25:1057-60.
53. Guyatt GH, Sullivan MJ, Thompson PJ, et al. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Can Med Assoc J* 1985;132:919-23.
54. McDonald CM, Henricson EK, Han JJ, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle & nerve* 2010;41:500-10.
55. Geiger R, Strasak A, Trembl B, et al. Six-minute walk test in children and adolescents. *J Pediatr* 2007;150:395-9, 9.e1-2.

56. Casey AF, Wang X, Osterling K. Test-retest reliability of the 6-minute walk test in individuals with Down syndrome. *Arch Phys Med Rehabil* 2012;93:2068-74.
57. Mayeux R. Biomarkers: Potential Uses and Limitations. *NeuroRx* 2004;1:182-8.
58. Pfeiffer G, Horvath R, Klopstock T, et al. New treatments for mitochondrial disease-no time to drop our standards. *Nature reviews Neurology* 2013;9:474-81.
59. Koene S, de Laat P, van Tienoven DH, et al. Serum FGF21 levels in adult m.3243A>G carriers: Clinical implications. *Neurology* 2014;6: 578.
60. Marcus KA, Barends M, Morava-Kozicz E, Feuth T, de Korte CL, Kapusta L. Early detection of myocardial dysfunction in children with mitochondrial disease: An ultrasound and two-dimensional strain echocardiography study. *Mitochondrion* 2011.
61. Bates MG, Hollingsworth KG, Newman JH, et al. Concentric hypertrophic remodelling and subendocardial dysfunction in mitochondrial DNA point mutation carriers. *Eur Heart J Cardio-vasc Imaging* 2013;14:650-8.
62. Hollingsworth KG, Gorman GS, Trenell MI, et al. Cardiomyopathy is common in patients with the mitochondrial DNA m.3243A>G mutation and correlates with mutation load. *Neuromuscul Disord* 2012;22:592-6.
63. Administration FaD. FDA Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
64. Pagliano E, Moroni I, Baranello G, et al. Outcome measures for Charcot-Marie-Tooth disease: clinical and neurofunctional assessment in children. *J Peripher Nerv Syst* 2011;16:237-42.
65. Beenakker EA, Maurits NM, Fock JM, Brouwer OF, van der Hoeven JH. Functional ability and muscle force in healthy children and ambulant Duchenne muscular dystrophy patients. *Eur J Paediatr Neurol* 2005;9:387-93.
66. Parreira SL, Resende MB, Zanoteli E, Carvalho MS, Marie SK, Reed UC. Comparison of motor strength and function in patients with Duchenne muscular dystrophy with or without steroid therapy. *Arq Neuropsiquiatr* 2010;68:683-8.
67. Abel MF, Damiano DL, Blanco JS, et al. Relationships among musculoskeletal impairments and functional health status in ambulatory cerebral palsy. *J Pediatr Orthop* 2003;23:535-41.
68. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science* 1977;198:679-84.
69. Holsbeeke L, Ketelaar M, Schoemaker MM, Gorter JW. Capacity, capability, and performance: different constructs or three of a kind? *Arch Phys Med Rehabil* 2009;90:849-55.
70. Bonnemann CG, Rutkowski A, Mercuri E, Muntoni F. 173rd ENMC International Workshop: congenital muscular dystrophy outcome measures 5-7 March 2010, Naarden, The Netherlands. *Neuromuscul Disord* 2011;21:513-22.
71. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27:S217-32.
72. Opstal SL, Jansen M, van Alfen N, de Groot IJ. Health-Related Quality of Life and Its Relation to Disease Severity in Boys With Duchenne Muscular Dystrophy:: Satisfied Boys, Worring Parents--A Case-Control Study. *Journal of child neurology* 2013;17:17.
73. Juniper EF, Price DB, Stampone PA, Creemers JP, Mol SJ, Fireman P. Clinically important improvements in asthma-specific quality of life, but no difference in conventional clinical indexes in patients changed from conventional beclomethasone dipropionate to approximately half the dose of extrafine beclomethasone dipropionate. *Chest* 2002;121:1824-32.

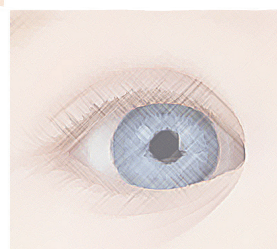
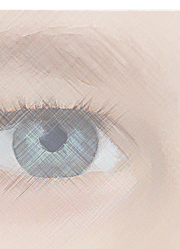
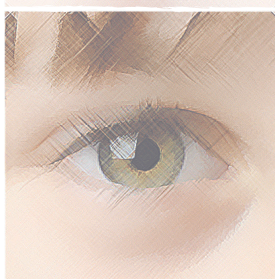
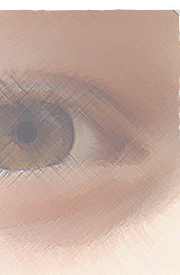
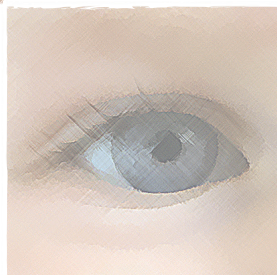
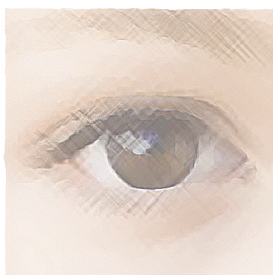
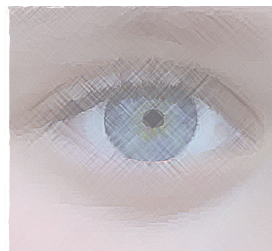
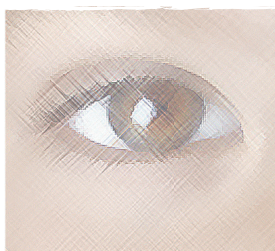
74. Davis SE, Hynan LS, Limbers CA, et al. The PedsQL in pediatric patients with Duchenne muscular dystrophy: feasibility, reliability, and validity of the Pediatric Quality of Life Inventory Neuromuscular Module and Generic Core Scales. *J Clin Neuromuscul Dis* 2010;11:97-109.
75. Hahn EA, Cella D, Chassany O, et al. Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clinic proceedings* 2007;82:1244-54.
76. Sullivan E, Barnes D, Linton JL, et al. Relationships among functional outcome measures used for assessing children with ambulatory CP. *Dev Med Child Neurol* 2007;49:338-44.
77. Oeffinger D, Gorton G, Bagley A, et al. Outcome assessments in children with cerebral palsy, part I: descriptive characteristics of GMFCS Levels I to III. *Developmental medicine and child neurology* 2007;49:172-80.
78. Battini R, Guzzetta A, Sgandurra G, et al. Scale for evaluation of movement disorders in the first three years of life. *Pediatr Neurol* 2009;40:258-64.
79. Battini R, Sgandurra G, Petacchi E, et al. Movement disorder-childhood rating scale: reliability and validity. *Pediatr Neurol* 2008;39:259-65.
80. Beghi E, Regio V, Papantonio A, et al. Reliability of clinical diagnosis of dystonia. *Neuroepidemiology* 2014;43:213-9.
81. Boyd RN, Graham HK. Objective measurement of clinical findings in the use of botulinum toxin type A for the management of children with cerebral palsy. *European Journal of Neurology* 1999:S23-S35.
82. Voet NB, van der Kooi EL, Riphagen, II, Lindeman E, van Engelen BG, Geurts A. Strength training and aerobic exercise training for muscle disease. *Cochrane Database Syst Rev*:CD003907.
83. McDowell BC, Kerr C, Parkes J, Cosgrove A. Validity of a 1 minute walk test for children with cerebral palsy. *Developmental medicine and child neurology* 2005;47:744-8.
84. Mayhew A, Mazzone ES, Eagle M, et al. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. *Developmental medicine and child neurology* 2013;55:1038-45.
85. Mazzone ES, Messina S, Vasco G, et al. Reliability of the North Star Ambulatory Assessment in a multicentric setting. *Neuromuscul Disord* 2009;19:458-61.
86. Huber AM, Feldman BM, Rennebohm RM, et al. Validation and clinical significance of the Childhood Myositis Assessment Scale for assessment of muscle function in the juvenile idiopathic inflammatory myopathies. *Arthritis and rheumatism* 2004;50:1595-603.
87. Omori CH, Silva CA, Sallum AM, et al. Exercise training in juvenile dermatomyositis. *Arthritis care & research* 2012;64:1186-94.
88. Elson JL, Cadogan M, Apabhaj S, et al. Initial development and validation of a mitochondrial disease quality of life scale. *Neuromuscul Disord* 2013;23:324-9.
89. Montes J, Gordon AM, Pandya S, De Vivo DC, Kaufmann P. Clinical outcome measures in spinal muscular atrophy. *Journal of child neurology* 2009;24:968-78.
90. Lynch DR, Farmer JM, Wilson RL, Balcer LJ. Performance measures in Friedreich ataxia: potential utility as clinical outcome tools. *Mov Disord* 2005;20:777-82.
91. Outline Measures Review and Selection Guidelines. (Accessed 14 July, 2015, at <http://www.researchchrom.com/reference/sr>.)
92. Mayhew JE, Florence JM, Mayhew TP, et al. Reliable surrogate outcome measures in multicentre clinical trials of Duchenne muscular dystrophy. *Muscle & nerve* 2007;35:36-42.
93. Filipovic Pierucci A, Mariotti C, Panzeri M, et al. Quantifiable evaluation of cerebellar signs in children. *Neurology* 2015;84:1225-32.



94. Spencer CT, Bryant RM, Day J, et al. Cardiac and clinical phenotype in Barth syndrome. *Pediatrics* 2006;118:e337-46.
95. Barth PG, Scholte HR, Berden JA, et al. An X-linked mitochondrial disease affecting cardiac muscle, skeletal muscle and neutrophil leucocytes. *J Neurol Sci* 1983;62:327-55.
96. Jefferies JL. Barth syndrome. *American journal of medical genetics Part C, Seminars in medical genetics* 2013;163C:198-205.
97. Rigaud C, Lebre AS, Touraine R, et al. Natural history of Barth syndrome: a national cohort study of 22 patients. *Orphanet journal of rare diseases* 2013;8:70.
98. Roberts AE, Nixon C, Steward CG, et al. The Barth Syndrome Registry: distinguishing disease characteristics and growth data from a longitudinal study. *American journal of medical genetics* 2012;158A:2726-32.
99. Yamashita S, Nishino I, Nonaka I, Goto Y. Genotype and phenotype analyses in 136 patients with single large-scale mitochondrial DNA deletions. *J Hum Genet* 2008;53:598-606.
100. Grady JP, Campbell G, Ratnaike T, et al. Disease progression in patients with single, large-scale mitochondrial DNA deletions. *Brain* 2014;137:323-34.
101. Onakpoya IJ, Spencer EA, Thompson MJ, Heneghan CJ. Effectiveness, safety and costs of orphan drugs: an evidence-based review. *BMJ open* 2015;5:e007199.
102. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. *Orphanet journal of rare diseases* 2008;3:11.
103. Cornu C, Kassai B, Fisch R, et al. Experimental designs for small randomised clinical trials: an algorithm for choice. *Orphanet journal of rare diseases* 2013;8:48.
104. Gupta D, Lammersfeld CA, Vashi PG, et al. Bioelectrical impedance phase angle in clinical practice: implications for prognosis in stage IIIB and IV non-small cell lung cancer. *BMC cancer* 2009;9:37.



# Summaries



# Summary

The aim of this thesis was to contribute to the knowledge on which outcome measures should be used in future clinical trials in children with mitochondrial disorders. This led to the following research goals: i) to explore which the symptoms and disabilities associated with mitochondrial disease should be measured in clinical trials from a patients' and from a doctors' perspective; ii) to systematically select instruments from literature that could be used to cover these symptoms and disabilities; iii) to validate these instruments in patients with mitochondrial disorders; and iv) to develop new instruments for domains, which seem important but are not covered. This thesis started with a general introduction into mitochondria, mitochondrial disease and mitochondrial medicine including the methodological challenges associated with the testing of potential medicines in this group of disorders.

### Studying interventions in mitochondrial disorders: past, present and future

**Part I** of this thesis focused on the methodological quality of pharmacological intervention studies performed in patients with mitochondrial disorders and the selection of clinically relevant and robust outcome measures for children with mitochondrial disease from literature.

**Chapter 1** critically reviewed the existing literature on treatment studies in mitochondrial disease. Although a large number of studies (1,039) were found, only 35 of them included more than five patients. When the methodological quality of these studies was rated, only 4 out of 35 studies had high-quality design. None of these high-quality studies had a positive outcome. Maybe even more worrying, many poor-design studies reported a positive influence of treatment on a non-predefined biomarker of unclear clinical significance, suggesting reporting of positive studies only. On the other hand, over the years, there was a promising trend towards higher-quality studies. Based on these results, recommendations for future studies in mitochondrial disease were formulated, including the use of validated, clinically meaningful and pre-specified primary end points and the development of biomarkers that are indicative of clinically relevant outcomes.

**Chapter 2** laid the groundwork for our subsequent studies. Using a questionnaire which was sent to all children seen in the Nijmegen Centre for Mitochondrial Disorders, we have made an inventory of the most burdensome complaints in of patients and their caretakers. In the 78 patients and caretakers who filled out the questionnaire, we found a high prevalence of tiredness and lack of energy, developmental delay, hypotonia, and balance problems. The most burdensome complaints reported by the cohort overall were tiredness, lack of energy, muscle weakness, speech and language problems, and epilepsy (as rated by children) as well as developmental delay and behavioural problems (mentioned by the parents). When the most burdensome complaints were rated only in the patients in which the symptom was present, epilepsy, hearing loss, depression, and

headache were amongst the most burdensome complaints for children. Strikingly, of the parents with a child with behavioural problems, 60% reported this symptom among their 3 most burdensome complaints, compared to 39% for epilepsy (rank number 2). When the treating physicians were asked which symptoms they estimated to be most burdensome to children with mitochondrial disorders and their parents, most paediatricians correctly mentioned fatigue, developmental delay, and muscle weakness, but the impact of e.g. speech and language problems and behavioural problems was underestimated. For this reason, we recommend to involve a sounding board of patients and/or caretakers in the development of clinical trial protocol design. After all, from a patient perspective, only they will know the feasibility of the protocol as well as the clinical relevance of the end point used.

The complaints and symptoms rated as most burdensome in **Chapter 2** were used in **Chapter 3** to select outcome measures. By systematic review of the published literature we made an inventory of the existing instruments. Subsequently, by using the experience in clinically similar disorders such as neuromuscular disorders and cerebral palsy, we indicated the psychometric properties of the available test. Finally, using strict criteria adopted by the Food and Drug Administration, we composed a toolbox consisting of 33 instruments with good face validity. Part of these instruments are clinically tested in Part III.

### Natural disease course

**Part II** of this thesis described the natural history of two groups of patients with mitochondrial disorders: those with a mutation in one of the nuclear genes leading to Complex I deficiency, and those with a mutation in mitochondrial DNA, the m.3243A>G mutation. One of the main goals was to make a rational choice on which population to include in the outcome measure validation studies.

In **Chapter 4**, the disease course of 130 patients with a nuclear encoded Complex I deficiency was described. Although these patients have some clinical features in common (they usually present within the first year of their lives with failure to thrive and severe neurological disease), the heterogeneity is striking. Many patients with Complex I deficiency and Leigh syndrome are described, but the course of the disease and the involvement of organs other than the brain is unpredictable and heterogeneous. Even within the same family, the clinical and biochemical phenotype may vary widely. This review showed that there is huge phenotypical heterogeneity within patients with the same genetic condition, in patients with the same biochemical condition and in patients with the same clinical syndrome. Therefore, we could not suggest a homogeneous population to perform our validation studies in.

**Chapter 5** studied 82 m.3243A>G carriers, a genetically homogeneous, but biochemically and clinically heterogeneous population. Although the m.3243A>G mutation was previously known as the MELAS (mitochondrial encephalopathy lactic acidosis and stroke-like episodes) mutation, we only detected three patients with MELAS syndrome in our cohort. The vast majority of the carriers (68%) suffered from Maternally inherited diabetes and deafness (MIDD), either or not in combination with muscle complaints or multi-organ involvement. This study again showed a huge clinical variability in m.3243A>G carriers. However, since such a high number of patients are available (we currently have ~150 carriers in our cohort), we have used this cohort for our surrogate marker validation studies.

## Validation studies

**Part III** aimed to test several of the outcome measures selected in **Chapter 3** in patients with mitochondrial disorders.

In **Chapter 6** we've tested the feasibility of the accelerometer, a device able to measure acceleration in a 3 dimensional plane. We asked 17 patients and 16 healthy controls to wear accelerometers on the upper and lower arm, the trunk and the upper leg during two weekend days at home. We found that children with a mitochondrial disease take more rest and have a lower peak intensity of movement, compared to age- and gender matched controls. When four patients and four healthy controls were tested again in a different season with different weather, similar results were obtained (good test-retest reliability). Since tiredness and lack of energy are among the most burdensome symptoms for patients and accelerometry is an objective and universally applicable measurement in a home instead of a laboratory situation, we think this is a promising and highly clinically relevant outcome measure for heterogeneous cohorts of children with mitochondrial disease. However, technical issues such as hardware failures, the labour intensity of the analyses and the correction for wheelchair movement need to be addressed before being applied in a clinical trial setting. Moreover, the study protocol should limit the external factors influencing the activity of the child during the measured weekend.

Since the scale which was available to rate general mitochondrial disease severity was not designed to monitor disease severity during intervention studies, we developed a new general Mitochondrial Disease Scale for children in **Chapter 7**. This score, the International Paediatric Mitochondrial Disease Scale (IPMDS), was designed and validated in collaboration with various international colleagues. The final questionnaire consisted of three domains: i) complaints and symptoms; ii) physical examination and iii) functional tests. The test was feasible, reliable and valid in a highly heterogeneous population of 17 children of various ages with mild, moderate and severe disease and functional disabilities from five centres around the world. In young children, the raters



experienced some difficulties in ‘translating’ the questions, but this did not hamper reliability. More studies have to be executed to know from which age the IPMDS is applicable. Because of the enormous heterogeneity in clinical phenotype, we expect that measures for general mitochondrial disease severity may be more generally applicable than symptom specific measures. Therefore, the IPMDS may be of use (as a secondary outcome measure) in future clinical trials. Although the inter-rater variability of the IPMDS is good, I would suggest depending on one or two well-trained raters only.

The 6-minute walking test (6MWT) is a widely used functional outcome measure for children with neuromuscular disease. In **Chapter 8** we’ve tested an alternative for this test, namely the (motor-)assisted 6-minute cycling test (A6MCT), in nine ambulatory children with OXPHOS dysfunction. Two children were not able to complete the test because of difficulties understanding the test, behavioural problems and/or attention deficit. For another two patients, face validity was low because the stability of the bicycle seemed to be the limiting factor in the number of revolutions made (due to a lack of coordination and relative high strength). Based on these results, we concluded that the A6MCT is not useful for ambulatory children with OXPHOS dysfunction. Because the test showed good face validity in children with more severe muscle weakness and the A6MCT with the legs correlate do the 6MWT, we still think this is a promising outcome for the small population of non-ambulatory children who are able to complete a 6-minute task.

**Chapter 9** assessed the feasibility and validity of 2-dimensional speckle tracking echocardiography (2DSTE) in 30 adult m.3243A>G carriers of which two subsequent echocardiograms were available. In 60 echocardiograms obtained as part of clinical care, the deformation of the myocardium was assessed using 2DSTE. Speckle tracking was feasible in 90% for longitudinal strain and for 57% in circumferential strain and radial strain. Circumferential and radial strain showed low face validity (low number of high quality images; suboptimal tracking) and were therefore rejected for further analysis. Global longitudinal strain had good face validity and was abnormal in 53 – 68% (depending on reference values used) of all carriers. Twenty-three (77%) patients could be included in the follow-up study. The difference between the first and the second global longitudinal strain measurement exceeded the measurement variance in 39% of the cases with time-intervals varying between 1.2 and 3.0 years. Based on these data, we concluded that 2DSTE is a promising outcome measure for future clinical studies and that 2DSTE should be tested in more detail in a prospective (intervention) study.

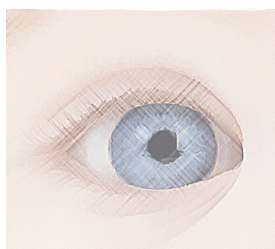
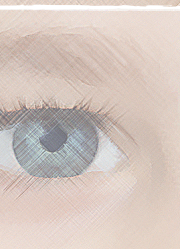
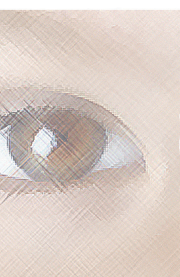
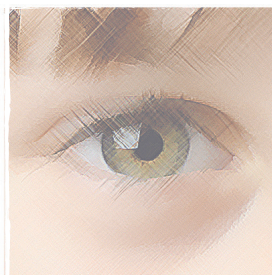
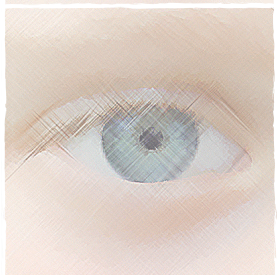
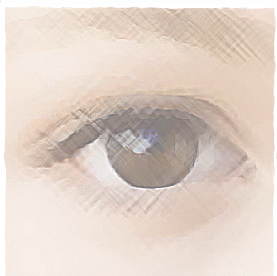
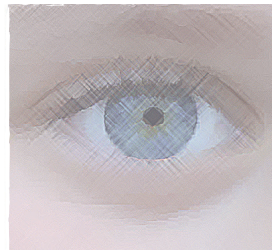
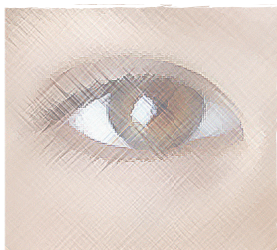
In **Chapter 10**, the value of fibroblast growth factor 21 (FGF21) as a surrogate measure for mitochondrial disease severity was tested in a cohort of 99 adult m.3243A>G carriers. Although there was a moderate correlation between FGF21 and general mitochondrial disease severity, a correlation between the change of FGF21 and the change in general mitochondrial disease severity over time lacked. Therefore, we decided that

this biomarker can not be used to monitor disease progression in m.3243A>G carriers. In **Chapter 11**, we came to the same conclusion for Growth and Differentiation Factor 15 (GDF15). Since GDF15 was previously described as a biomarker for cardiac disease, we also evaluated the correlation between the concentration of GDF15 and myocardial deformation (measured by the same method as described in **Chapter 9**). Although GDF15 cannot be used as a biomarker for general mitochondrial disease severity, the moderate correlation between the concentration of GDF15 and myocardial deformation might suggest that it can be used as a biomarker for mitochondrial cardiomyopathy.

### Final remarks

This thesis concluded with a general discussion including future perspectives. Although we cannot draw solid conclusions on which end points to use in future clinical trials, I make some cautious recommendations on which outcome measures to use in future clinical trials as well as on clinical trial design. First of all, I recommend to confirm the efficacy of a drug in adult patients with mitochondrial disease, before a high-quality study evaluating the efficacy, safety and pharmacokinetics in the paediatric population is initiated. For the paediatric trial, I would recommend either a randomised, placebo-controlled, double-blinded Bayesian design in all children with mitochondrial disease or a randomised, placebo-controlled, double-blinded cross-over study in a more homogeneous and clinically stable cohort of children. As end points for clinical trials in paediatric mitochondrial disease, I recommend further optimisation and evaluation of the accelerometer, the IPMDS, the Gross Motor Function Measure (GMFM), 2DSTE, the Pediatric Evaluation of Disability Inventory (PEDI), the 6MWT, and the timed function tests. These studies are currently ongoing.



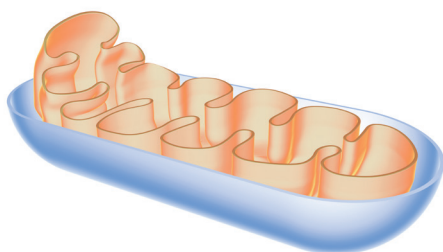


# Samenvatting

In deze samenvatting geef ik eerst een korte inleiding over mitochondriële ziekten en waarom het zo noodzakelijk is een medicijn voor deze aandoeningen te ontwikkelen. Vervolgens zal ik toelichten welke uitdagingen zich voordoen als het effect van een potentieel medicijn getest wordt. Ten slotte zal ik toelichten wat dit proefschrift bijdraagt aan onze kennis met betrekking tot het testen van potentiële medicijnen in deze groep patiënten.

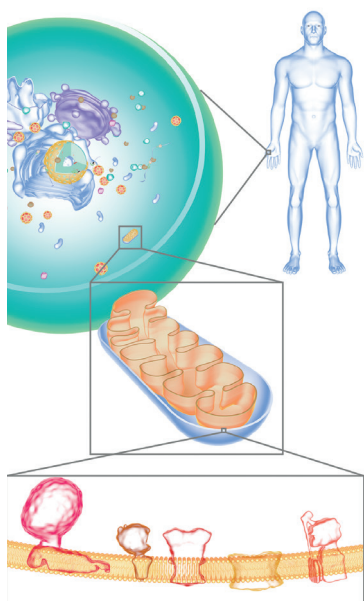
Mitochondriële ziekten worden ook wel energiestofwisselingsziekten genoemd. Voor bijna alle processen die in ons lichaam plaatsvinden is energie nodig. Deze energie wordt aangeleverd via voedingsstoffen zoals koolhydraten, vetten en eiwitten en door de cellen omgezet in chemische energie waar de cel wat mee kan. Deze chemische energie wordt ook wel adenosine trifosfaat (ATP) genoemd. ATP wordt gemaakt in de energiefabrieken van de cel, de mitochondriën. ATP wordt niet alleen gebruikt om te bewegen en te functioneren, maar ook om de cellen in ons lichaam te onderhouden. Een mens zet ongeveer zijn eigen lichaamsgewicht in ATP om per dag! Mitochondriën zijn onderdelen van de cel, naast bijvoorbeeld de celkern en de eiwit producerende organellen. De mitochondriën leveren de energie voor de andere celonderdelen om goed te kunnen functioneren. Het aantal mitochondriën per cel varieert, afhankelijk van de energiebehoefte van de cel: de hersenen, het hart, de lever en de skeletspieren bevatten veel mitochondriën, terwijl de huid er veel minder bevat. De meeste lichaamscellen bevatten ongeveer 500 tot 2000 mitochondriën.

Een mitochondrion bestaat uit twee membranen die door een ruimte zijn gescheiden (Figuur 1). Rondom deze membranen kunnen drie compartimenten onderscheiden worden: de binnenkant van het mitochondrion (matrix), de buitenkant van het mitochondrion (de vloeistof in de cel, het cytosol of cytoplasma) en ruimte tussen de membranen (de tussen-membraan ruimte). Het binnenmembraan heeft uitstulpingen naar binnen waardoor het sterk geplooid is. Hierdoor heeft het een groot oppervlak, wat van groot belang is voor de energieproductie die hier plaats vindt.



**Figuur 1.**  
Een artist impression van een mitochondrion met een dubbele membraan, waarvan de binnenste geplooid is.

Het produceren van energie is een ingewikkeld proces dat begint met het afbreken van voedingsstoffen zoals koolhydraten, vetten en eiwitten in kleine stukjes. Tijdens de afbraak wordt acetyl-coA geproduceerd, waaraan in de citroenzuurcyclus energierijke elektronen worden onttrokken. Deze elektronen worden door een vijftal grote eiwitcomplexen (Complex 1 tot 5 (I-V); Figuur 2) op de mitochondriële binnenmembraan omgezet in ATP.



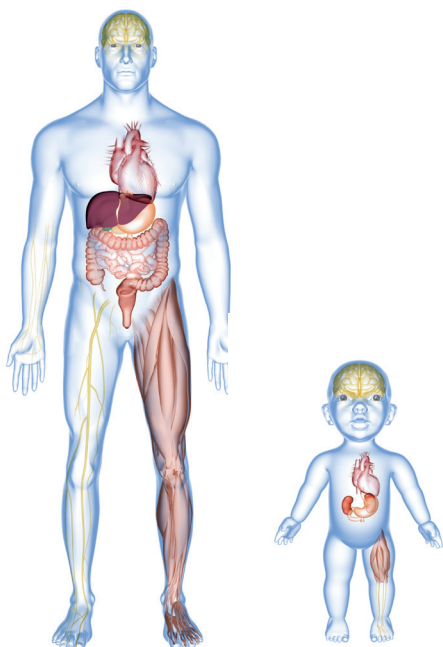
**Figuur 2.**

**In een cel zitten vele mitochondriën.** Op de binnenmembraan van de mitochondriën zitten de eiwitcomplexen die energie maken.

Lichaamscellen hebben de in de mitochondriën geproduceerde energie nodig om goed te kunnen functioneren. Als een van de eiwitcomplexen niet goed werkt ontstaat er een energietekort, waardoor cellen en organen onvoldoende kunnen functioneren. Meestal wordt het niet goed functioneren van een eiwitcomplex veroorzaakt door een genetische afwijking. Deze worden meestal óf van beide ouders geërfd (autosomaal recessief) óf alleen van de moeder (mitochondriële overerving; tijdens de bevruchting worden de mitochondriën van de vader verwijderd). De eerstgenoemde DNA foutjes verdelen zich gelijk over het hele lichaam. De DNA fouten van de moeder kunnen in wisselende percentages (heteroplasmie) per cel aanwezig zijn.

Patiënten met een verminderd vermogen om energie te produceren, een mitochondriële ziekte, hebben meestal klachten van de organen die het meest energie verbruiken,

zoals de hersenen, het hart, de ogen, de lever en de skeletspieren (Figuur 3). De uitingvormen van deze ziekten kunnen extreem verschillend (heterogeen) zijn, niet alleen tussen patiënten, maar ook binnen één patiënt gedurende zijn leven. Aankomend artsen wordt geleerd dat ze moeten denken aan een mitochondriële aandoening bij 'elk onverklaard symptoom, van elk orgaan, op elke leeftijd, met iedere vorm van overerving'. De oorzaak voor deze grote verschillen is nog niet helemaal duidelijk.



**Figuur 3.**

**Overzicht van de organen die het meest frequent zijn aangedaan bij patiënten met mitochondriële ziekten.** Van boven naar beneden: de hersenen, de ogen, het hart, de lever, de maag en darmen, de spieren en de zenuwen.

Kinderen zijn vaak het meest ernstig aangedaan. Sommige kinderen overlijden als ze pas geboren zijn, anderen zijn ernstig geestelijk en lichamelijk gehandicapt. Er zijn ook kinderen die normaal presteren op school, maar zich slecht in kunnen spannen. De ziekte kan zich ook op latere leeftijd openbaren en heeft dan meestal een milder beloop.

Er is op dit moment nog geen behandeling voor deze patiënten, maar er zijn wel een groot aantal studies in patiënten uitgevoerd. Echter, als je deze studies op een rijtje zet (Hoofdstuk 1), blijken slechts 12 van de meer dan 1000 studies op een wetenschappelijk acceptabele manier te zijn uitgevoerd, waarvan er slechts 4 aan de hoogste standaarden voldoen! De meeste studies voldeden niet aan de juiste methodologie (geran-



domiseerde gecontroleerde studies), hadden een hoge mate van bevooroordeeldheid (bias) of waren maar in een handvol patiënten uitgevoerd.

Een belangrijk aspect van het goed uitvoeren van een klinische studie is het selecteren van uitkomstmaten. Uitkomstmaten zijn de tests die de onderzoekers uitvoeren om het effect van de behandeling te beoordelen. Voorbeelden van uitkomstmaten zijn vragenlijsten, bloedtesten, conditietesten of beeldvorming. Het is belangrijk dat deze uitkomstmaten betrouwbaar en valide zijn. Immers, als je uitkomstmaat onafhankelijk van het ziekteproces verschillende waarden aanneemt, is het lastig te bepalen welk effect bepaald werd door het geteste medicijn.

In dit proefschrift heb ik op een wetenschappelijke en systematische manier bestudeerd welke uitkomstmaten geschikt zouden kunnen zijn voor het meten van het effect van therapie bij kinderen met een mitochondriële ziekte. Hiervoor ging ik eerst na welke klachten patiënten ervaren en welke klachten zij als meest belangrijk beschouwen (Hoofdstuk 2). De meest belastende klachten bleken vermoeidheid, gedragsstoornissen, moeite met spreken, epilepsie en spierzwakte. Ook rapporteerden ouders dat hun kinderen erg veel beperking in het functioneren in het dagelijks leven ervaren. Een opvallende bevinding van deze studie was dat als de behandelend artsen gevraagd werden welke klachten zij dachten dat ouders en patiënten het meest belastend zouden vinden, dit niet volledig overeen kwam met de klachten die als meest belastend werden gerapporteerd.

Naast een inventarisatie van de gerapporteerde klachten van de kinderen die wij vervolgen, werd ook het natuurlijk verloop van de klachten in twee verschillende groepen patiënten onderzocht. In patiënten met een mutatie in een van de kern genen die zorgen voor een verminderde werking van het eerste complex van de oxidatieve fosforylering, zagen wij dat patiënten ontzettend van elkaar verschillen (Hoofdstuk 4), en dat zelfs binnen families met dezelfde genetisch afwijking totaal verschillende klachten kunnen voorkomen. Wij concluderen in dit overzichtsartikel dat er niet alleen binnen patiënten met hetzelfde genetische en biochemische defect, maar ook binnen patiënten met hetzelfde klinische syndroom een enorme variabiliteit is. Het gebruiken van een van deze indelingen om vergelijkbare groepen samen te stellen is dus van weinig belang. Daarom kozen wij ervoor om alle kinderen met mitochondriële aandoeningen mee te laten doen aan onze validatie studies.

Ook bij patiënten met de meest voorkomende mutatie in het mitochondriële DNA (de m.3243A>G mutatie) zagen wij een enorme variatie in betrokken organen en ernst van de klachten en beperkingen (Hoofdstuk 5). Echter, aangezien het aantal dragers van de m.3243A>G mutatie zo groot was, leek dit toch een van de meest geschikte populaties om onze studies in te doen.

Met in ons achterhoofd de klachten die voor patiënten het meest belangrijk waren en daarnaast de klachten die in studies in het natuurlijk beloop veel bleken voor te komen, keek ik welke instrumenten er op dat moment beschikbaar waren voor het meten van de eerder gevonden klachten (Hoofdstuk 3). Hierbij heb ik me met name gericht op die instrumenten die een klacht meten die relevant is voor de patiënt, zoals uithoudingsvermogen, motorische capaciteiten, zelfverzorging en kwaliteit van leven. Omdat er binnen het onderzoek naar mitochondriële ziekten nog weinig aandacht besteed was aan dit onderwerp, hebben wij hierbij gekeken naar de resultaten in andere ziekten zoals Duchenne spierdystrofie of aangeboren hersenbeschadiging.

De hierboven geselecteerde uitkomstmaten, zoals een ondersteunde fietstest en een bewegingsmeter, heb ik vervolgens getest in patiënten met mitochondriële ziekten (Hoofdstuk 6-11). Hierbij keek ik naar de haalbaarheid in de beoogde patiëntenpopulatie, de betrouwbaarheid en de relevantie van de test voor de patiënt.

Wij vroegen 17 kinderen met een mitochondriële ziekte tijdens een weekend een aantal bewegingsdetectoren te dragen (Hoofdstuk 6). Voor de meeste kinderen was het haalbaar de meters gedurende het hele weekend te dragen. Vergeleken met de 16 gezonde kinderen van dezelfde leeftijd en hetzelfde geslacht die wij gemeten hebben, bewogen kinderen met mitochondriële ziekten minder, hadden zij minder hoge pieken in hun bewegingspatroon en namen zij meer rust. Voordat de bewegingsdetectoren gebruikt kunnen worden in toekomstige studies, moeten er echter nog wel een aantal technische problemen worden opgelost. Daarnaast moeten externe factoren die het bewegingspatroon van het kind beïnvloeden tijdens het meten zoveel mogelijk worden uitgeschakeld.

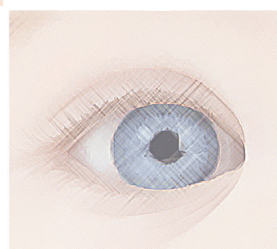
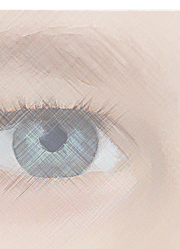
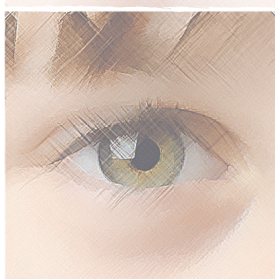
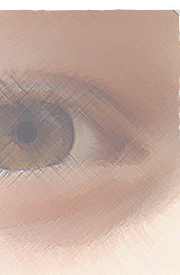
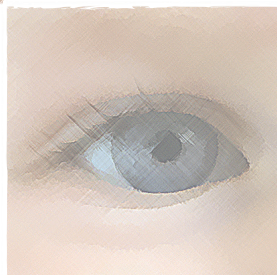
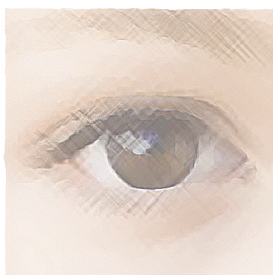
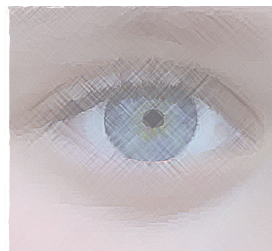
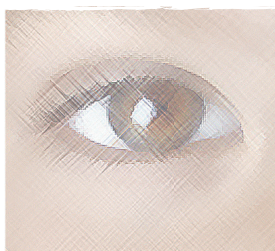
Omdat er geen maat is om de algemene ziekte-ernst van kinderen met een mitochondriële ziekte te vervolgen in studies waarbij het effect van een medicijn getest wordt, ontwikkelden wij hiervoor een meetinstrument (Hoofdstuk 7). Deze vragenlijst, de IPMDS, werd ontwikkeld en getest in samenwerking met internationale collega's en bestaat uit 3 onderdelen: de klachten van de patiënt, het lichamelijk onderzoek en de functionele mogelijkheden van de patiënt. Bij de studie door 15 internationale collega's bleek dat de IPMDS haalbaar, betrouwbaar en valide leek in een heterogene groep met 17 kinderen van verschillende leeftijden en met verschillende aandoeningen en maten van beperking. Ook jonge kinderen konden betrouwbaar gemeten worden, al moeten we nog wel onderzoeken vanaf welke leeftijd de IPMDS goed lijkt te meten. Aangezien er weinig testen zijn die bij alle kinderen haalbaar zijn en een klinisch relevante beperking of symptoom meten, lijkt algemene ziekte-ernst een veelbelovende uitkomstmaat. Gezien de goede meeteigenschappen van de IPMDS kan deze goed voor dat doel gebruikt worden.

Ik testte ook een ondersteunde fietstest, waarbij kinderen gevraagd werden om gedurende 6 minuten zo veel mogelijk rondjes te maken met hun armen of benen (Hoofdstuk 8). Hieruit bleek dat als de kinderen nog goed in staat waren om 6 minuten te lopen, deze test minder geschikt leek. Ook was de test niet haalbaar in kinderen met een ernstige ontwikkelingsachterstand, gedragsproblemen of concentratieproblemen. Aangezien er maar weinig kinderen zijn die te zwak zijn om 6 minuten te lopen, maar zich wel gedurende 6 minuten in kunnen zetten voor een fietstest, heeft het testen van de ondersteunde fietstest niet de hoogste prioriteit.

Ten slotte deed ik twee onderzoeken in volwassen dragers van de m.3243A>G mutatie. Wij analyseerden hartecho's met een techniek waarbij de vervorming van de hartspier uitgedrukt kan worden in een getal (Hoofdstuk 9). Hieruit bleek dat meer dan de helft van de dragers een afwijkende vervorming van de hartspier had. Toen wij de vervorming van de hartspier na enige tijd weer hebben gemeten, leken wij ook in staat veranderingen op te merken. Dit zal echter nog bevestigd moeten worden in grotere studies. Ik testte bij deze dragers ook twee biomarkers (Hoofdstuk 10 en 11), stoffen die in het bloed gemeten kunnen worden en waarvan gedacht werd dat zij iets zeggen over de aanwezigheid en de ernst van de ziekte. Aangezien de spreiding van de concentratie van de biomarkers niet veel leek te zeggen over de ernst van de ziekte, concludeerde ik dat deze nog niet gebruikt kunnen worden als uitkomstmaat.

Ik sluit af met een discussie van de onderzoeksresultaten, waarbij ik ook naar de toekomst kijk. Hierin doe ik aanbevelingen voor het gebruik van uitkomstmaten in toekomstige studies en schuif ik voorzichtig een aantal tests naar voren waarvan ik denk dat ze het verloop van de ziekte het best in kaart zullen brengen. Daarnaast doe ik de aanbeveling om medicijnen in volwassenen te testen, alvorens het effect van deze medicatie op kinderen te bepalen.

Met het onderzoek gepresenteerd in dit proefschrift hoop ik een aanzet gegeven te hebben tot het verbeteren van de kwaliteit van klinische studies in kinderen met een mitochondriële ziekte. Door het onderzoeken van de uitkomstmaten voordat er een studie uitgevoerd wordt, hopen we te voorkomen dat studies worden uitgevoerd waarvan van tevoren al gezegd had kunnen worden dat de resultaten niet waardevol zullen zijn.



Dankwoord



## Dankwoord

Het zit erop! Een mooie en dynamische periode waarin ik veel geleerd heb en veel interessante en leuke mensen ontmoet heb. Ik wil iedereen dan ook bedanken voor deze mooie tijd, zowel degenen met wie ik samen heb mogen werken, als degenen die me daarbuiten weer de energie en motivatie gaven om er weer vol tegenaan te gaan.

Prof. Dr. J.A.M. Smeitink, Beste Jan. Om kansen te grijpen moet je ze wel eerst krijgen. Ik ben dankbaar voor de zichtbare lol die je eruit haalt om de jongere generatie kansen te geven en te laten groeien. Bedankt ook voor het vertrouwen dat ik mijn originele AGIKO-project idee een eigen richting mocht geven en voor de vrijheid m'n ideeën uit te werken tot projecten. Ook in minder plezierige periodes was je flexibel en dacht je mee over nieuwe kansen, welke uiteindelijk denk ik goed gaan uitpakken. Je bevologenheid heeft me echt gestimuleerd om ook het uiterste uit mezelf te halen! Ik hoop dat je binnenkort een medicijn kunt afleveren dat wel werkt.

Dr. I.J.M. de Groot, Beste Imelda, wat fijn dat jij mijn co-promotor was! Jouw combinatie van klinisch inzicht en contacten met bewegingswetenschappers hebben onder andere tot ons mooie accelerometer project geleid. Maar ook je kritische commentaren op protocollen en manuscripten heb ik erg gewaardeerd; je wist altijd de vinger op de zere plek te leggen, wat de kwaliteit van de stukken echt ten goede is gekomen. Ik hoop dat we nog vele jaren kunnen samenwerken!

Leden van de manuscriptcommissie, Geachte prof. dr. F.G.M. Russel en prof. dr. J.P.H. Drenth, hartelijk dank dat ik mijn proefschrift ter beoordeling aan u mocht voorleggen. Dear professor P.F. Chinnery, thank you for your time spent in judging my thesis and the effort of coming to Nijmegen for my defence (and King's Day).

Ook wil ik de patiënten bedanken die vrijwillig hun spaarzame energie investeerden voor de medische wetenschap: bedankt jongens en meiden dat jullie mee wilden doen aan de fietstest, de IPMDS en het onderzoek met de bewegingsdetector! Ik hoop dat we snel een medicijn vinden dat jullie meer energie zal geven om te spelen, te leven en jullie toekomstdromen waar te maken. Beste ouders, dank ook voor de tijd dus jullie wilden investeren in wéér een ziekenhuisbezoek. Nu ik zelf moeder ben van een kleintje met bijna onbepaalde energie kan ik me iets beter voorstellen hoe het zou moeten zijn als hij dat niet had. Nog een extra reden om hard door te werken aan een medicijn! Ook de deelnemers aan de nationale inventarisatie van m.3243A>G dragers wil ik hartelijk bedanken voor hun tijd en moeite.

Ik wil ook mijn collega's bedanken voor de prettige samenwerking, waarvan sommige mensen in het bijzonder. Mirian, met je betrokkenheid bij zowel individuele patiënten als

de opzet van de klinische trial ben je voor mij echt hét voorbeeld van een academische arts. Lonneke en Maaïke, dank voor jullie bijdrage aan de accelerometer studie en de IPMDS pilot. Lotte en Lonneke, hopelijk wordt de LHON studie net zo'n succes als de m.3243A>G studie! Floor, leuk dat je mee wilde doen met het instructiefilmpje van de IPMDS! Marion, we leerden elkaar pas echt kennen in het gammele stapelbedje in Stockholm. Fantastisch hoe jij de PCS studie en de coördinatie van de MitoRoute op je hebt genomen! Annette, dank voor je vliegensvlugge antwoorden en met het versturen van de 143 brieven voor Hoofdstuk 2. Gera, wat goed dat we jouw versterking hebben aan het Mito-front.

Professor Morava, beste Eva, zonder jouw aanstekelijke enthousiasme was ik nu zeker niet geweest waar ik nu ben. Miski, in 2003 liepen we samen intro, nu gaan we in hetzelfde veld promoveren. De lunches met jou en Thatjana, waren altijd even een momentje van ontspanning (en bovendien was ik dan ook weer op de hoogte van alle relevante roddels). Succes met de afronding van jullie proefschriften en opleidingen. Paul, de betrokkenheid waarmee jij het m.3243A>G cohort door de jaren heen gezien hebt zou wel eens het verschil zou kunnen gaan maken tijdens onze eerste trial. Heel veel succes met het afronden van je opleiding en proefschrift. Wing and Sheila, thanks for your help with the IPMDS during your traineeship in Nijmegen. Ilse, wat ben ik blij dat ik je bij het begin van je derde collegejaar een sms'je gestuurd heb. Niet alleen om je hulp en gezelligheid tijdens de fietstesten, IPMDS studie en accelerometer pilot, maar ook omdat ik blij ben getuige te zijn geweest van de groei die je hebt doorgemaakt als arts, onderzoeker en als mens. Esmeë, wat een zeldzaam goede onderzoeksstage heb je neergezet! Heel veel succes met je co-schappen, je wordt vast een fantastische dokter. Robert, je inzet getuigt van een ambitieniveau waar je ver mee gaat komen. Heel veel plezier in the Big Apple. Pascal, al ben je nog geen jaar betrokken bij het onderzoek, je voelt als volwaardig onderdeel van ons team.

Dear co-authors, thank you very much for your contribution to the articles included in this thesis! Merel, dankjewel voor je hulp bij het corrigeren van ons review toen ik in Zuid-Amerika was, daardoor heb ik onbezorgd van m'n sabbatical kunnen genieten. Leuk ook om 'jouw' fietstest in onze populatie te proberen! Snel een keer samen thee doen? Arjen, Kenneth en Hans, jullie technische/Matlab ondersteuning op gelegen en minder gelegen momenten was onmisbaar voor het accelerometer project. Doorlène, ik heb ongeloofelijk veel bewondering hoe je tussen je behandelingen door het biomarker onderzoek toch in eigen hand hebt weten te houden. Anjo, fijn dat je de patiënten voor de accelerometer studie accelerometer studie voor ons wilde beoordelen. Richard, dank voor het geduldig antwoorden van al mijn mailtjes over lijsten met patiënten (gelukkig wordt het nu makkelijker gemaakt dit vanuit de kliniek zelf bij te houden). Livia, dankjewel voor je begeleiding tijdens het exploreren van de voor mij totaal nieuwe wereld van de 2DSTE. Gert, respect voor je engelengeduld met de vele (extra) analyses



voor onze 2DSTE studie. Je speedcursus Access die ook in mijn andere onderzoeken goed van pas is gekomen! Ik wil alle leden van het MitoRoute team bedanken voor hun enthousiaste inzet tijdens de zorg, maar daarmee ook de dataverzameling voor toekomstig uitkomstmaten-onderzoek. Ook de mensen van het spierlab, bedankt voor het analyseren van de dubbele spierbipten voor de studie die helaas dit proefschrift niet gehaald heeft. Thanks also to all members the Radboud Center for Mitochondrial Medicine, although I couldn't always fully understand, your blots and complexomes learned me a lot about basic mitochondrial medicine.

Kees en Maroeska, dank voor jullie vertrouwen dat ik aan de opleiding tot klinisch farmacoloog kon beginnen, ondanks dat ik niet aan het standaard profiel voldoe. Charlotte, Wim, Simone, Marysia, Adi, Sofie, Jenneke, Marjolein en de andere 'opleidingelingen', dank voor de gezelligheid en interactie op de dinsdag. Dat maakt de opleiding levendig! Kinderartsen uit het CWZ, bedankt voor de tijd dat ik klinische ervaring bij jullie op heb mogen doen. Het was wat korter dan gepland, maar ik heb veel geleerd, over de praktijk, maar vooral ook over mezelf.

Niet alleen de mensen in het ziekenhuis verdienen een dankwoordje, misschien juist de mensen buiten het ziekenhuis die mijn blikversmalling overleefd hebben.

Union Dames 5, na een dag achter de computer was het altijd heerlijk om achter een balletje aan te kunnen hollen en eventuele frustraties weg te slaan. Bedankt voor de ontspanning en de gezelligheid elke dinsdag en zondag. Suze, Dien, An en Ing, echt super bijzonder dat we al sinds 2003 samen spelen en nu ook nog 'lotgenoten' in het combineren van het moederschap en alle andere ballen die hoog gehouden moeten worden. Die baby-date moet er nu echt maar es komen!

Myrtille, Bianca en Inge, gezellig om samen alle fasen tijdens en na de co-schappen te doorlopen. Lieve Myrtille, ik ben ongelofelijk trots dat je doorgezet hebt en toch doctor bent geworden. Je wordt vast een fantastische gynaecoloog! Gezellig dat jullie nu om de hoek wonen, daar moeten we vaker gebruik van maken. Lieve Inge, leuk om zo veel levensfasen samen te doorlopen. Succes met de afronding van jouw boekje! Beste Bianca, we moeten snel weer een theetje doen!

Beste Marlies en Koen, dankjulliewel voor onze ontspannen bacchanaaltjes met veel vinologisch en culinair plezier. Beste Dirk, Claudia en Emme, dankjulliewel voor de ontspannen koppels-met-kind middagen. Na Bolivia en München nu met z'n vieren moeten we binnenkort maar eens wat met z'n achten ondernemen!

Laura en John, ondanks de afstand voelen jullie toch altijd dichtbij. Lieve Lau, je hebt al m'n dipjes en euforische momenten van dichtbij mee mogen maken. Ook als we elkaar

even niet zien, ben je er voor mijn gevoel altijd. Ik zou niet weten wat ik zonder je zou moeten. Dankjewel dat je ook nu weer aan mijn zijde staat. Beste John, dankjewel dat je zo goed voor m'n beste vriendinnetje zorgt. Binnenkort weer Zeeland, Kaatsheuvel of Disneyland? Lieve Joris en Sophia, wat een vrolijk stel zijn jullie! Ik gun jullie als jullie groot zijn een vriendje of vriendinnetje die net zo lief is als jullie moeder (en vader natuurlijk).

Lieve Marij en Harrie, dank voor al jullie steun de afgelopen jaren. De weekenden in Tilburg en later ook in Villa Gladiola Nova waren altijd een welkome afwisseling op het drukke leven op het werk. Harrie, zonder jou was het niet mogelijk geweest om in ons paleisje te wonen, waar we de rust vinden om hard te kunnen werken. Marij, hoe fijn is het om een tweede moeder te hebben die er zichtbaar van geniet om voor ons gezinnetje te zorgen! Een ding is zeker: dit boekwerk was er nog niet geweest zonder jullie hulp!

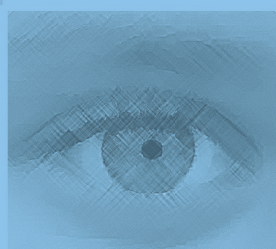
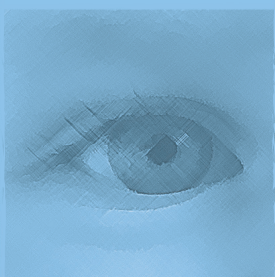
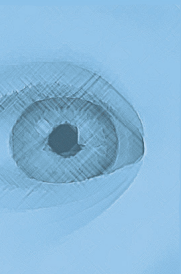
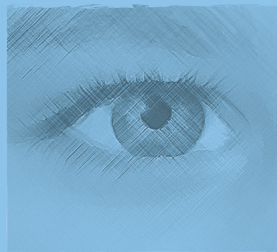
Erik en Nienke, jullie hebben heel wat te verduren gehad met die ingewikkelde gesprekken over geneeskunde tijdens de kerstdiners. Broertje, ook al zien we elkaar niet zo vaak, we weten dat we er op de belangrijke momenten voor elkaar zijn. Zo ook vandaag weer. Dankjewel dat je mijn paranimf wil zijn! Ik hoop dat we er nog veel mooie momenten mogen delen. Schoonzusje, bedankt dat je mijn kleine broertje gelukkig maakt! Tebbe, lieve rouwdouwer, woef!

Pap en mam, jullie waren altijd geïnteresseerd in waar ik mee bezig was, ook al werd het soms een beetje veel jargon en afkortingen in huis als Dennis en ik langs kwamen. Pap, ondanks dat je me altijd ontraden hebt dokter te worden, denk ik dat – via wat omwegen – ik toch m'n plek en uitdaging gevonden heb. Dat eigenwijze karakter heb ik ook niet van een vreemde, hè? Mam, dankjewel voor je interesse in mijn wel en wee en de ontspannen momentjes, eerst in de stad en nu bij ons thuis aan de eettafel. Pap, mam, dit proefschrift is de kroon op jullie opvoeding waarin ik altijd mocht ontdekken en zelf mocht doen en waar nieuwsgierigheid, doorzettingsvermogen en een kritische houding gestimuleerd werden. Hopelijk kan ik de volgende generatie Vriensjes ook de autonome opvoeding geven die ik van jullie kreeg.

Dennis, Lief Aapje, dank voor je hulp als co-auteur en statistiekleraar, al was dat niet je belangrijkste bijdrage. Ik ben echt gegroeid door je stimulans het uiterste uit mezelf te halen en buiten mijn comfortzone te treden. Met de Argentijnse tango heb je me helemaal bewezen dat dingen die je niet kunt ook heel leuk kunnen zijn! Bovendien: waar vind je iemand die het ook een goed idee vindt om 5 dagen op een boot te zitten om pinguïns en orka's in het wild te zien, zwanger-en-wel midden in de nacht af te dalen in de krater van een actieve vulkaan, 5 uur om te rijden om tenminste samen in de file te

staan en om met een baby van 4 maanden rond te reizen in Zuid-Afrika? Er is een hoop veranderd de afgelopen jaren...en ik zou nooit meer anders willen!

Lieve Thijntje, jouw nieuwsgierigheid, ondernemingsdrang, vrolijkheid en energie geven elke dag kleur. Wat ben ik blij dat jij er bent!



# List Of Publications and Curriculum Vitae

## Peer Reviewed Original Articles

**Koene S**, Timmermans J, Weijers G, de Laat P, de Korte CL, Smeitink JAM, Janssen MCH, Kapusta L. Is 2D speckle tracking echocardiography useful in detecting and monitoring myocardial dysfunction in adult m.3243A>G carriers? – a retrospective pilot study. Submitted

**Koene S**, Hendriks J, Dirks I, de Boer L, de Vries MC, Janssen MCH, Smuts I, Fung CW, Wong VCN, de Coo IFM, Vill K, Klopstock T, Falk M, McCormick EM, McFarland R, Hendriks JCM, de Groot IJM, Smeitink JAM. The International Paediatric Mitochondrial Disease Scale. Submitted

**Koene S**, Dirks I, van Mierlo E, de Vries PR, Janssen AJWM, Hendriks JCM, Smeitink JAM, Bergsma A, Essers HJMN, Meijer K, de Groot IJM. The value of using 3D accelerometry in estimating daily physical activity in children with mitochondrial disease. Submitted

Verhaak CM, de Laat P, **Koene S**, Tibosch M, Rodenburg RJT, Knoop H, Janssen MCH, Smeitink J. Quality of life, fatigue and mental health in patients with the m.3243A>G mutation correlates with genetic characteristics and disease manifestation. Orphanet Journal of Rare Diseases, in press

Wong SSN, Goraj B, Fung CW, Vister J, de Boer L, **Koene S**, Smeitink J. The Nijmegen Centre for Mitochondrial Disorders Paediatric Mitochondrial Disorder MRI Score. Submitted

Dirks I, **Koene S**, Verbruggen R, Smeitink JAM, Jansen M, de Groot IJM. The assisted 6-minute cycling test: an exploratory study in children. Conditionally accepted for publication by Muscle & Nerve

**Koene S**, de Laat P, van Tienoven DH, Vriens D, Sweep FCGJ, Timmersmans J, Kapusta L, Janssen MCH, Smeitink JAM. Serum GDF15 levels correlate to mitochondrial disease severity and myocardial strain, but not to disease progression in adult m.3243A>G carriers. JIMD Reports Epub May 13 2015

Dirks I, **Koene S**, Verbruggen R, Smeitink JAM, Jansen M, de Groot IJM. An Exploring study of the psychometric properties of the Assisted 6-Minute Cycling Test in Children with Neuromuscular and Mitochondrial Disorders. Submitted for publication

**Koene S**, de Laat P, van Tienoven DH, Vriens D, Sweep FCGJ, Timmersmans J, Kapusta L, Janssen MCH, Smeitink JAM. Serum GDF15 levels correlate to mitochondrial disease severity and myocardial strain, but not to disease progression in adult m.3243A>G carriers. *JIMD Reports* 2015;24:69-81.

**Koene S**, de Laat P, Vriens D, van Tienoven DH, Brandt AM, Sweep FCGJ, Borm G, Janssen MCH, Smeitink JAM. Serum Fibroblast Growth Factor 21 (FGF-21) levels in adult m.3243A>G carriers: clinical implications. *Neurology*. 2014 Jul 8;83(2):125-33. Pfeffer G, Horvath R, Klopstock T, Mootha VK, Suomalainen A, **Koene S**, Hirano M, Zeviani M, Blindoff LA, Wai Man PY, Hanna M, Varelli V, McFarland R, Majamaa K, Turnbull DM, Smeitink J, Chinnery PF. New treatments for mitochondrial disease: no time to drop our standards. *Nat Rev Neurol*. 2013 Aug;9(8):474-81.

**Koene S**, Wortmann SB, de Vries MC, Jonckheere AI, Morava E, de Groot IJM, Smeitink JAM. Developing outcome measures for paediatric mitochondrial disorders: which complaints and limitations are most burdensome to patients and their parents? *Mitochondrion*, 2013 Jan;13(1):15-24.

de Laat P, **Koene S**, vd Heuvel LPWJ, Rodenburg RJT, Janssen MCH, Smeitink JAM. Inheritance of the m.3243A>G mutation. *J Inherit Metab Dis Reports*, 2013;8:47-50.

de Laat P, **Koene S**, vd Heuvel LPWJ, Rodenburg RJT, Janssen MCH, Smeitink JAM. Clinical features and heteroplasmy in blood, urine and saliva in 34 Dutch families carrying the m.3243A>G mutation. *J Inherit Metab Dis* 2012 Nov;35(6):1059-69.

Morava E, Gardeitchik T, Kozicz TL, de Boer L, **Koene S**, de Vries MC, McFarland R, Roobol T, Rodenburg RJT, Verhaak CM. Depressive behaviour in children diagnosed with mitochondrial disorder, *Mitochondrion* 2010 Aug;10(5):528-33

## Peer Reviewed Review Articles

Koopman WJH, Beyrath J, Fung CW, **Koene S**, Rodenburg R, Willems PHGM, Smeitink JAM. Pathophysiology of OXPHOS-related disorders in children: Towards development of clinically relevant treatment strategies. *EMBO Mol. Med*, in press.

Fung CW, **Koene S**, Willems, PHGM, Koopman WJH, Smeitink JAM. Emerging roles of mitochondrial translation, coenzyme Q biosynthesis, membrane homeostasis and quality control in mitochondrial oxidative phosphorylation: clinical phenotypes, diagnostic signatures and treatment strategies. Submitted

**Koene S**, Jansen M, Verhaak CM, de Vroeh RLA, de Groot, IJM, Smeitink JAM. Towards harmonisation of Outcome Measures in Children with Mitochondrial Disorders. *Dev Med Child Neurol*. 2013 Aug;55(8):698-706.

**Koene S**, Rodenburg RJ, van der Knaap MS, Willemsen MAAP, Sperl W, Laugel V, Ostergaard E, Tarnopolsky M, Martin MA, Nesbitt V, Fletcher J, Edvardson S, Procaccio V, van den Heuvel LPWJ, Smeitink JAM. Natural disease course and genotype-phenotype correlations in Complex I deficiency: what we learned from the more than 130 known cases. *J Inherit Metab Dis*. 2012 Sep;35(5):737-47

**Koene S**, Smeitink JAM. Metabolic manipulation. *J Inherit Metab Dis*. 2011 Apr;34(2):315-25

Breuer ME, Koopman WJ, **Koene S**, Nooteboom M, Rodenburg, RJT, Willems, PJ, Smeitink JAM. The role of OXPHOS dysfunction in the development of neurologic disease. *Neurobiology of Disease* 2013 Mar;51:27-34.

**Koene S**, Willems PHGM, Roestenberg P, Koopman WJH, Smeitink JAM. Mouse models for mitochondrial complex I deficiency, *J Inherit Metab Dis*. 2011 Apr;34(2):293-307

**Koene S**, Smeitink JAM. Mitochondrial medicine, entering the era of treatment, *J Intern Med*. 2009; 265(2):193-209.

## Peer Reviewed Case Reports and Case Series

**Koene S**, Matthyssens L, Buzing C. Een zuigeling met twee huiddefecten ter plaatse van de kleine fontanel. *Praktische Paediatric*; June 2013; 92,

**Koene S**, Rodenburg R, Peters G, de Groot I, Verhaak C, Kapusta L, Smeitink JAM, Morava E. Muscle pain, fatigue and hypothermia at night in association with mitochondrial dysfunction, *Journal of Paediatric Neurology*, 2009, 1304-2580

**Koene, S**, Kozicz, TL, Rodenburg RJT, Verhaak C, de Vries MC, van den Heuvel LPWJ, Smeitink JAM, Morava E. Major depression in adolescent children consecutively diagnosed with mitochondrial disorder, *J Affect Disord*.;114(1-3):327-32. 2009

**Koene, S**, Kluijtmans L, Wevers R, Pasch M, Mock D, Morava E. Biotin-sensitive 3-methylcrotonylglycinuria in a patient with severe growth delay, ectodermal abnormalities, neonatal progeroid appearance and developmental delay. *Clin Dysmorphol*.; 17(3):195-8 2008



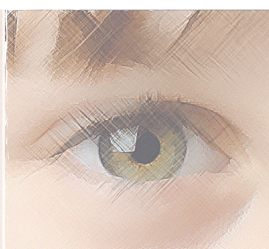
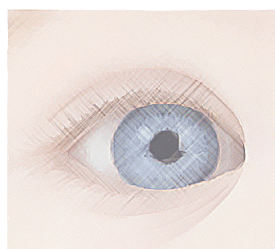
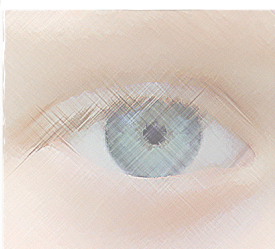
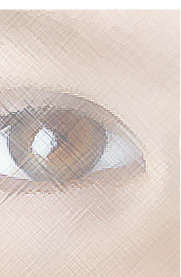
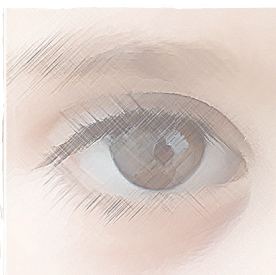
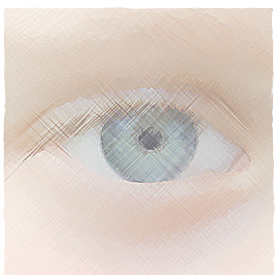
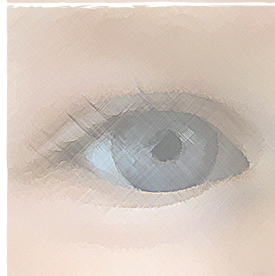
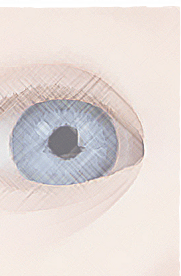
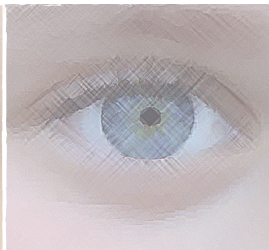
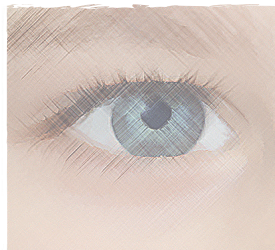
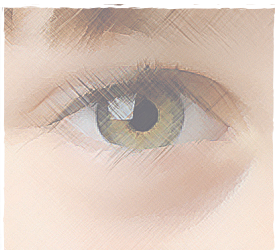
## Other Publications

**Koene S**, Smeitink J. Mitochondrial Medicine. Editorial. J Inherit Metab Dis. 2011 Apr;34(2):247-8

Vogel RO, **Koene S**, Smeitink JAM. Complex I. Book Chapter in Scriver, the Online Metabolic and Genetic Basis of Inherited disease. <http://www.ommbid.com>; McGraw Hill Publications.

## Books

Mitochondrial Medicine – a clinical guideline, **Saskia Koene** and Jan Smeitink; ISBN 978.90.81.773. First press 2011, second press 2013.



# Curriculum Vitae



## Curriculum Vitae

Saskia Koene was born on the 19th of May 1984 in Weert, the Netherlands. She grew up in Wageningen. After her graduation, she started her medical training at the Radboud University in Nijmegen. During her medical training she was involved in various committees of the Medical Faculty Students Association Nijmegen (MFVN) and the Nijmegen Students Hockey Association (NSHC) Apeliotes, was quaestor of the 11th lustrum of the MFVN (2007) and was the praeses of NSHC Apeliotes (2004 – 2005).

Saskia started her research career as a student-researcher under supervision of Eva Morava, MD, PhD in 2007. They wrote her first two papers together and set up the excellent student trajectory, in which young students get the opportunity to participate in scientific research. After one year, Jan Smeitink, MD, PhD, MAE, took on the supervision and they laid the foundations for her PhD project and published three literature reviews during Saskia's internships. Her final internship was at Great Ormond Street Hospital in London under supervision of P.E. Clayton, MD, PhD and A. Male, MD, PhD.

In 2010, she started a PhD project on the treatment of mitochondrial diseases at the Nijmegen Centre for Mitochondrial Disorders, still under supervision of Jan Smeitink. Her PhD trajectory was funded with an Qualified Doctor Training to become a Clinical Researcher (AGIKO) grant by Netherlands Organisation for Health Research and Development (ZonMW). When the focus on outcome measures became more clear, Imelda de Groot, MD, PhD agreed to be a co-supervisor. For her studies, Saskia received financial support from the Tjallingh Roorda foundation, Stofwisselkracht and the Rare Disease Foundation. She was invited to present her work in several lectures on national and international conferences. Saskia was also part of the National Institute of Health (NIH) working group on common data elements for mitochondrial disease. To create more awareness and understanding of mitochondrial diseases, she wrote the booklet 'Mitochondrial Medicine - a clinical guideline' together with Jan Smeitink. During her PhD trajectory, she also followed 9 months of clinical training in paediatrics in the Canisius Wilhelmina hospital under supervision of Ben Semmekrot, MD, PhD.

Since January 2015, she has started her training in clinical pharmacology with a focus on orphan drugs trial design, under supervision of Kees Kramers, MD, PhD and Maroeska te Loo, MD, PhD. Besides, she is continuing her research on outcome measures in mitochondrial disease.

Saskia resides with Dennis in Nijmegen and they have a son, Thijn (12-10-2014). They're expecting their second child in July 2016. She likes to travel the world with her family and to play field hockey.